PHENOTYPIC AND GENOMIC ASSESSMENT OF LISTERIA MONOCYTOGENES VIRULENCE

A Dissertation
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Maria Ximena Cardenas Alvarez

In Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

Major Program:
Molecular Pathogenesis

July 2019

Fargo, North Dakota

North Dakota State University
Graduate School

Title

PHENOTYPIC AND GENOMIC ASSESSMENT OF LISTERIA MONOCYTOGENES VIRULENCE

By

Maria Ximena Cardenas Alvarez

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

DOCTOR OF PHILOSOPHY

SUPERVISORY COMMITTEE:

Teresa M. Bergholz
Chair

Paul Carson

John McEvoy

Mark Strand

Approved:

09/30/19 — Date

John McEvoy — Department Chair

ABSTRACT

Listeria monocytogenes is the etiological agent of listeriosis in humans and ruminants causing bacteremia, central nervous system (CNS) infections, abortion, and gastroenteritis among other clinical outcomes. Recent studies have integrated whole genome sequence (WGS), epidemiology data, and host susceptibility to provide evidence for variation in virulence among strains, as a small number of hypervirulent clones have been found linked to a high proportion of human and ruminant invasive listeriosis cases, however, still little is known about variation in virulence across different *L. monocytogenes* subgroups.

To assess and compare the genetic diversity of clinical listeriosis isolates from ruminants in the Upper Great Plains states, we used multilocus sequence typing (MLST) and found that the variation in virulence potential varies among clonal complex (CC), which is reflected in the epidemiology of *L. monocytogenes.* Based on these results, we evaluated the strains' virulence potential in *Galleria mellonella* through larvae survival, LD_{50}, and cytotoxicity, and monitored health index scores and bacterial concentrations post-infection as quantifiable indicators of virulence and immunogenicity. Our findings suggest that strains belonging to CC14, as well as isolates from MN infections are hypervirulent in *G. mellonella*, as they need a lower bacterial concentration to cause disease and produce a low-level infection that could help in evading the host immune response. We also identify genomic elements associated with strains causing three different clinical outcomes: bacteremia, central nervous system infections, and maternal-neonatal infections. By analyzing 232 whole genome sequences from invasive listeriosis cases, we identified orthologous genes of phage phiX174, transfer RNAs and type I restriction-modification (RM) system genes along with SNPs in loci associated with environmental

adaptation such as *rpoB* and the phosphotransferase system (PTS) associated with one or more clinical outcomes.

Novel genetic variants may be associated with a particular virulence phenotype, as it is likely that strains causing the same clinical outcome share unique genetic elements. Variation in virulence among *L. monocytogenes* subgroups may confer an increased ability to cross host barriers or higher adaptability to food processing environments, thus the investigation of strain-specific genetic features can impact the design of prevention and management plans for listeriosis.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor Dr. Teresa Bergholz, for the continuous support during these years. Her patience, motivation, and commitment have been an inspiration for my career. Likewise, I would like to thank my committee: Dr. Paul Carson, Dr. John McEvoy, and Dr. Mark Strand for their insightful comments and encouragement, as well as Dr. Peter Bergholz for his valuable feedback and support with the computer cluster.

My special thanks to the Francisco Jose de Caldas Scholarship for Doctoral Studies from COLCIENCIAS (Colombia) for their financial support through these years, and to the Department of Microbiological Sciences at NDSU for giving me the opportunity to participate as Graduate Teaching Assistant.

DEDICATION

Simon

Daniel

Mom & Dad

Sebastian

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDIX TABLES

LIST OF APPENDIX FIGURES

1. LITERATURE REVIEW

1.1. *Listeria monocytogenes* as human and ruminant pathogen

The *Listeria* genus includes 17 bacteria species divided into two groups: *Listeria* sensu stricto, which includes *L. monocytogenes*, *L. seelgeri*, *L. ivanovii*, *L.marthii*, *L. welshimeri*, and *L. innocua*; and *Listeria* sensu lato, which consists of another 11 species. Within the genus, only *L. monocytogenes* and *L. ivanovii* are considered pathogenic (Bakker et al., 2014; Orsi and Wiedmann, 2016). *L. monocytogenes* is a ubiquitous Gram-positive facultative intracellular bacillus that causes disease in humans and animals. It has been defined as a zoonotic disease transmitted between animals and humans as a consequence of direct contact, indirect environmental contact, or through food (European Food Safety Authority (EFSA)European Centre for Disease Prevention and Control (ECDC), 2017). *L. monocytogenes* exists in the environment as a saprotroph and can access the human food chain either directly or through infection or carriage in farm animals (Gahan and Hill, 2014). It is frequently found as a contaminant in food and animal feed (Dhama et al., 2015) causing listeriosis, a foodborne infectious disease acquired through oral route that affects specific population groups such as immunocompromised hosts, elderly, pregnant women and newborns (Lomonaco et al., 2015) and causes different clinical manifestations in humans and farmed ruminants such as central nervous system (CNS) infections, bacteremia (BAC), stillbirth, late-term spontaneous abortions and mastitis (Nightingale et al., 2005).

In farmed ruminants, listeriosis has a big impact in economic losses in livestock production due to the morbidity and high mortality in animals as well as on food safety and public health since affected animals and healthy carriers (up to 50%) may spread the bacteria contaminating produce, water and milk products (Borucki et al., 2004; Zundel and Bernard,

2006). Throughout the farm environment, *L. monocytogenes* has been detected in 8-22% of water trough or tanks, 11% of bedding samples and 8-37% of soil samples, highlighting the potential for spread of the pathogen (Dreyer et al., 2016; Garcia et al., 1996). Thus, livestock and produce are considered the primary sources for the introduction of *L. monocytogenes* into the food chain (Fenlon et al., 1996; Sauders et al., 2012).

Variance in virulence among strains of *L. monocytogenes* has been observed both *in vivo* in animal models as well as through epidemiological data. Epidemiological data indicates that not all strains of *L. monocytogenes* are equally capable of causing disease as isolates from only four (out of 13) serotypes identified (1/2a, 1/2c, 1/2b, 4b) are responsible for more than 98% of the human listeriosis reported and are distributed mostly among lineage I and II (Orsi et al., 2011). To date *L. monocytogenes* isolates represent four phylogenetic lineages based on the differences in the nucleotide sequence of specific genes, as in multilocus sequence typing (MLST) and multilocus virulent genes sequence (MVLST), or differences in the whole genome (wgMLST) or core genome (cgMLST) (Datta and Burall, 2018). Lineage I have been associated with human clinical cases and outbreaks mainly caused by serotypes 1/2b and 4b, and has been described as clonal due to a low prevalence of plasmids and insertion sequence (IS) elements that suggests a limited acquisition of foreign DNA by mechanisms of recombination; lineage II groups strains from food, environment, and animal listeriosis cases mostly caused by serotype 1/2a and 1/2c and shows a higher recombination rate that possibly facilitates adaptation to diverse environments. Lineage III and IV are very diverse with recombination levels between those for lineage I and II and include mostly isolates obtained from ruminants, however, there are few isolates analyzed to date (Chlebicz and Śliżewska, 2018; Orsi et al., 2011).

L. monocytogenes lineages can be subdivided in multiple clonal complexes (CC), which are groups of genetically related isolates that are presumed to have descended from a recent common ancestor and are assigned using genotypic approaches, such as MLST (Wiedmann, 2002). Epidemiological and microbiological data collected in the last few years indicate that the strength of the association of a CC with a clinical disease could be linked to its gene content. In 2016, Maury *et al* showed that the most prevalent CCs differ based on source: CC1, CC2, CC4, and CC6 were strongly associated with a clinical origin, whereas CC121 and CC9 were strongly associated with a food origin. Likewise, using mice as an animal model, three categories of highly prevalent clones were distinguished: infection-associated clones (CC1, CC2, CC4, and CC6), food-associated clones (CC9 and CC121) and intermediate clones (CC3, CC5, CC8+CC16, CC37, CC155) (Maury et al., 2016a). These results suggest that there is evidence on how the presence of clonal groups seems to be characterized by unique virulence or host specificity patterns, however, it is not clear yet, why some strains are more likely to cause infections while others have been established as non-clinical/ food-associated strains.

1.2. Epidemiology of listeriosis

According to the World Health Organization (WHO), in 2010 around 600 million cases of diseases caused by contaminated food were reported, including 350 million caused by pathogenic bacteria. Bacterial diseases of animal origin caused by *Campylobacter* sp., *Salmonella* sp., *Listeria* sp., or the Enterobacteriaceae family, constitute a serious health risk both in developing countries and in advanced ones as well (Chlebicz and Śliżewska, 2018). Between 2005 and 2015, *L. monocytogenes* caused more foodborne outbreaks in the European Union (EU) (83) than in the United States (US) (47), resulting in 757 and 491 cases, respectively (European Food Safety Authority (EFSA)European Centre for Disease Prevention

and Control (ECDC), 2017), however, a higher number of cases requiring hospitalization (428 vs. 332) and more deceases (82 vs. 61) were reported in the US (Rodríguez-López et al., 2018). From 2016 to date, nine outbreaks have been reported in the US, related mainly to vegetables, raw dairy and deli meats, where 51 individuals were ill, 49 were hospitalized and 7 of them died (Dewey-Mattia et al., 2018).

Compared to other food- and feed-borne infections, listeriosis is not a common disease; still, the majority of its cases are associated with necessary hospitalization and treatment. The disease is also associated with a high mortality rate, reaching 20–30%, and for risk-group patients even 75% (Chlebicz and Śliżewska, 2018). The CDC's 2011 Estimates of Foodborne Illness, which includes estimates of cases not reported, indicates that around 1,600 cases and 255 deaths occur annually in the US, with an annual incidence of 0.24 cases per 100,000 population (Scallan et al., 2011) . Likewise, the Foodborne Diseases Active Surveillance Network (FoodNet), which covers about 15% of the US population, reported 158 cases and an incidence of infection per 100,000 population of 0.3 in 2017, 26% higher when compared with the incidence of listeriosis during 2014-2016. Additionally, hospitalization (94%) and mortality (15.9%) rates are ones of the highest when compared with other pathogens transmitted commonly by food, ranking as #1 and #3 respectively (Marder et al., 2018).

L. monocytogenes has a high and particularly severe incidence in elderly, immunocompromised individuals, pregnant women, and neonates. In these hosts, the invasive form of the illness can be symptomatically manifested as bacteremia and neurolisteriosis, or abortions and neonatal infections. Host risk factors for bacteremia and neurolisteriosis include older age, innate and cellular immune defects, malignancies, HIV infection, cirrhosis, diabetes mellitus, and immunosuppressive therapies, however, other features such as alcoholism, antacid

uptake, corticosteroid medication, chronic lung or kidney disease, and blood transfusion have been also recognized as risk factors for death due to *L. monocytogenes* (Charlier et al., 2017).

In the US, individuals aged more than 85 years are 54 times more likely to acquire listeriosis (95% CI: 37–79) compared to the 15–44-year-old population. Similarly, pregnant women were found to have a relative risk of listeriosis of 115 (95% confidence interval [CI] 69–205) compared to women of the same age (A. M. Pohl et al., 2017). A prospective clinical study focusing on all forms of invasive listeriosis in France found that more than 80% of infected mothers experienced major fetal or neonatal complications due to *L. monocytogenes* (fetal loss, very high prematurity, early or late onset disease); and only 39% of patients with neurolisteriosis survived and fully recovered. Likewise, they identified independent factors associated with mortality in bacteremia and neurolisteriosis such as ongoing cancer, multi-organ failure, decompensated comorbidity, monocytopenia, and concomitant bacteremia for neurolisteriosis, which were added to the list of risk factors identified previously (Charlier et al., 2017).

While there are many factors that have been associated with variation in the incidence of listeriosis, demographic changes in the US population must be considered as well. As the US population ages and the proportion of the Hispanic population grows (higher relative risk related to diet/socioeconomic factors), an increase in listeriosis cases and incidence rates would be expected if exposure and infectivity remain unchanged. The Healthy People 2020 program from the Office of Disease Prevention and Health Promotion (ODPHP) under their food safety objective looks for the reduction of foodborne illnesses in the US by improving food safety-related behaviors and practices. Their target for reducing the incidence rate of laboratory-confirmed cases of *L. monocytogenes* is 0.2 cases per 100,000 people, however, the incidence rates in the US have not shown a significant decrease in the last few years (Marder et al., 2018).

Infection with *L. monocytogenes* in ruminants has been linked to the ingestion of poorly fermented silage during the colder months; yet, this could also be related to the animal density and husbandry types used during winter (Nightingale et al., 2005). It has also been shown that bedding, water tanks, and barn equipment may be other significant sources of contamination apart from silage to explain increased fecal shedding and higher number of listeriosis cases during these months, however, the comparison of affected and unaffected ruminants and the study of the farm environment would contribute to the identification of infection sources and risk factors to control and reduce ruminant listeriosis (Steckler et al., 2018; Walland et al., 2015).

Preventing the entry of *L. monocytogenes* into the food chain is challenging due to its ubiquity and high-stress tolerance, allowing it to survive and persist under numerous environmental conditions (Freitag et al., 2009). *L. monocytogenes* circulates between animals, humans and food facility environments, affecting not only the health aspect but also the economy due to the cost of hospitalization, financial losses associated with food recalls, work absence, and legal proceedings. The implementation of prevention and management measures to limit the incidence of listeriosis is essential in order to assess and establish different control points (Chlebicz and Śliżewska, 2018; Walland et al., 2015).

1.3. Pathogenesis of *L. monocytogenes*

The infectious process of *L. monocytogenes* starts after the consumption of contaminated food. After crossing the intestinal barrier the bacterium reaches the liver and the spleen through the lymph and blood. There, it multiplies and disseminates via bloodstream reaching target organs like the brain and the placenta (Cossart and Toledo-Arana, 2008). Intracellular pathogens such as *L. monocytogenes* can cross several host barriers and enter the cytosol of a variety of cell types. It has the capacity to enter, survive and multiply in both phagocytic and non-phagocytic

cells, thus this property is considered to be essential for the pathophysiology of listeriosis. There are several factors that mediate host cell infection, including internalins responsible for adherence and internalization of bacteria into nonphagocytic host cells, listeriolysin O (LLO) a pore-forming cytolysin that acts together with phospholipases and allows escape from the phagocytic vacuole, and ActA that mediates actin-based cell to cell spreading (Lecuit, 2005).

Internalization uses a "zipper mechanism" that involves the progressive interaction of bacterial surface proteins with their specific cell receptors. InlA and InlB are the best-characterized internalins from a family of around 25 proteins. Respectively, they bind to E-cadherin (Ecad) and Met -the receptor of the hepatocyte growth factor (HGF), located on the surface of host cells and induce bacterial uptake through receptor-mediated endocytosis. This first step requires actin nucleation and polymerization, which leads to cytoskeletal rearrangement. After entry, the bacterial cell is trapped into a vacuole from where it escapes into the cytoplasm by using a pore-forming virulence factor LLO along with two phospholipases (PlcA and PlcB). Likewise, lipoprotein A (PplA), a peptide pheromone-encoding lipoprotein, is secreted altering the signaling and secretion of proteins involved in quorum sensing to mimic a larger bacterial population and contributing to the vacuolar escape (Lamond and Freitag, 2018; Radoshevich and Cossart, 2018). Once in the cytoplasm, the bacterium multiplies rapidly. ActA, a bacterial surface protein, induces polymerization of host cell actin filaments forming a network, which enable *L. monocytogenes* to migrate to the host cell margin and into cell-wall projections. These pseudopodia are engulfed and ingested by neighboring cells, in which the life cycle begins again without triggering the host's immunological response. Therefore, *L. monocytogenes* can infect cells by two different mechanisms: direct invasion and cell-to-cell spread (Chlebicz and Śliżewska, 2018; Cossart and Toledo-Arana, 2008).

1.4. Clinical manifestations of listeriosis

L. monocytogenes has the remarkable ability to cross three significant barriers: the intestinal barrier, the blood-brain barrier and the feto-placental barrier (Lecuit, 2005). When the infection is not invasive (listeriosis), febrile gastroenteritis is the predominant symptom of infection, characterized by affecting healthy hosts rather than immunocompromised individuals (Gahan and Hill, 2005). On the other hand, invasive listeriosis affects mainly immunocompromised hosts causing the most common forms of the infection: neurolisteriosis, maternal-neonatal infection, and bacteremia.

1.4.1. Central nervous system (CNS) listeriosis

Non-exclusive ways for bacteria to cross the blood-brain barriers have been described. Extracellular bacteria, either free in the blood and/or associated with cells, may recognize receptors at the surface of the barriers and cross them, or alternatively, the bacterium may gain the CNS in infected cells, such as circulating leucocytes that are known to be able to cross themselves the blood-brain barriers (Disson and Lecuit, 2012).

Meningitis and meningoencephalitis are the most common presentations of neurolisteriosis or CNS listeriosis in humans (70 to 97%) (Disson and Lecuit, 2012). Patients with *L. monocytogenes* meningitis experience a longer prodromal phase and present with signs and symptoms similar to those reported in the general population with community-acquired bacterial meningitis (Pagliano et al., 2017). Clinical characteristics are associated with a sub-acute course that lasts around 24 hours before admission in hospital and includes abnormal movements, seizures, and alteration of consciousness caused by the dissemination of the bacterium to the functional tissue in the brain (parenchyma). Known predisposing factors are

immunosuppression, age over 50 years old and underlying conditions such as malignancy or diabetes (Brouwer et al., 2006).

Contrary to meningitis and meningoencephalitis, rhombencephalitis occurs mostly in healthy individuals. Clinical signs of rhombencephalitis include headache, malaise, nausea and vomiting that can last 4- 10 days followed from more severe signs such as asymmetrical cranial nerve deficits and cerebellar signs such as inability to perform and sustain a series of rapidly alternating muscle movements (dysdiadokinesia and dysmetria), gross incoordination of muscle movements (ataxia), repetitive, involuntary oscillation of the eyes (nystagmus), wide tremor during voluntary movements (intention tremor), slow and distorted speech (slurred speech), and muscle weakness (hypotonia) (Disson and Lecuit, 2012).

In ruminants, the most common presentation of CNS listeriosis is rhombencephalitis (Walland et al., 2015). Its incubation period may vary from 1-7 weeks and some of the first signs observed include disorientation, depression, anorexia, head tilt and circling (from there it is also known as 'circling disease'). Facial paralysis is usually presented and affects only one side of the face, causing the ear and eyelid to droop. Profuse salivation may also be observed due to a lack of muscle tone in the lip of the affected side. Sick animals may also be unable to stand or fall when standing and also exhibit involuntary running movements while lying down (Oevermann et al., 2010). In cattle, disease progression is slower, has a longer overall duration and lower mortality compared with infections in small ruminants, which often die within 1 to 3 days after symptoms onset (Walland et al., 2015).

Two major routes of access of *L. monocytogenes* to the CNS have been proposed. While in humans *L. monocytogenes* passes the gastrointestinal barrier and spreads through the blood to the brain, in ruminants it is suggested that *L. monocytogenes* enters the cranial nerves via the oral

epithelium or conjunctivae, reaching the brain stem and spreading to other regions via axonal pathways without showing signs of systemic infection. This probably is related to the fact that *L. monocytogenes* induces almost exclusively rhombencephalitis in these animals (Disson and Lecuit, 2012; Walland et al., 2015).

1.4.2. Maternal-neonatal (MN) listeriosis

Maternal infection can present as asymptomatic or flu-like symptoms 1- 14 days prior to the appearance of fetal distress, making diagnosis difficult and resulting in late diagnosis and adverse outcomes for the fetus (Madjunkov et al., 2017). Neonatal listeriosis can happen as either: early onset that occurs during the first 7 days after birth and is associated with transplacental infection; or as late onset, which appears after the first week and could be related to transplacental infection, exposure during delivery or an external source. Listeriosis during pregnancy causes spontaneous abortion, premature birth, stillbirth and neonatal complications such as sepsis and meningitis (Elinav et al., 2014). Although *L. monocytogenes* can cause infection any time during pregnancy, it is most often diagnosed in the third trimester. Recent studies using nonhuman primate models have provided evidence that *L. monocytogenes* tropism for the maternal reproductive tract may result in infection of the decidua, placenta and the fetus from the first trimester of pregnancy (Wolfe et al., 2017).

Pregnant women have an estimated 18-fold increased incidence of disease compared with non-pregnant healthy individuals (Bakardjiev et al., 2005). Likewise, more than 80% of infected mothers experience major fetal or neonatal complications, and most of the fetal losses occur at less than 29 weeks of gestation and within 2 days of hospital admission (Charlier et al., 2017). Pregnancy-related listeriosis has been more commonly found among ethnic minorities in France, England and Wales, Hispanic pregnant women in the United States and LOTE (language other

than English) speakers in Australia, and is associated with consumption of contaminated food, lower socioeconomic conditions and unawareness of risks (Charlier et al., 2017; Crim et al., 2014; Dalton et al., 2011; Mook et al., 2010; Silk et al., 2012).

The immunological environment of the MN interface is unique since protection from infections and tolerance of the fetus by the maternal immune system has to be balanced. Opposite to the hypothesis that pregnant women are immunocompromised, recent studies have suggested that the maternal immune system is carefully regulated (Faralla et al., 2016). Interestingly, like *L. monocytogenes*, other human pathogens with mucosal portal of entry and at least partial intracellular life cycles, such as *Toxoplasma gondii*, *Treponema pallidum* and enteroviruses have in common not only to disseminate to the CNS but successfully target and multiply within the cells of the placenta and fetus, causing devastating effects (Lamond and Freitag, 2018). CNS infection does not occur in pregnant individuals, which suggests that pregnancy is not a predisposing factor for neurolisteriosis and more importantly that mechanisms for MN and CNS infections are different (Disson and Lecuit, 2012).

In general, the placenta consists of both maternal and fetal-derived cells, although cellular architecture varies among mammals. Human placenta is hemochorial, meaning that the maternal blood is in direct contact with the trophoblast cells of the embryo. It consists of branching villi structures, which includes both floating villi and villi anchored into the decidua formed by extravillous cytotrophoblast cells (EVT). It also includes a continuous layer of fused multinucleated syncytiotrophoblasts (SYN) that have differentiated from the underlying cytotrophoblast (CTB) cells and form the syncytia. *L. monocytogenes* can invade the placenta via direct invasion of SYN (in contact with maternal blood), or through cell to cell spread infecting EVT cells anchored into the decidua or from bacteria within the maternal leukocytes (Lamond

and Freitag, 2018). Pregnant ruminants have a different type of placentation. Instead of having a single large area of contact between fetal and maternal vascular systems, these animals have many smaller placentomes, and six layers of tissue separating the maternal and fetal bloodstreams (epitheliochorial). *L. monocytogenes* can localize in the placentomes, where maternal-fetal exchanges take place and enter the amniotic fluid reaching the fetus and causing different outcomes mainly in the late-term of gestation (Šteingolde et al., 2014).

1.4.3. Bacteremia

L. monocytogenes can be isolated from blood cultures without signs of neurolisteriosis or maternal-neonatal infection. Patients with bacteremia report influenza-like symptoms such as fever, tachycardia, and diarrhea making the diagnosis difficult in the population of patients at the highest risk, however, bacteremia is more likely to be diagnosed in patients with underlying conditions due to the constant monitoring through blood cultures (Hernandez-Milian and Payeras-Cifre, 2014). Multi-organ failure and decompensated comorbidity are more severe symptoms presented by patients with bacteremia, and incubation period might range between 1-10 days being 2 days the average (Chlebicz and Śliżewska, 2018). Invasive bacteremia is a common manifestation of listeriosis in the elderly, and in most cases, an origin of infection cannot be determined. Studies have shown that the increase in cases in this population is related to the frequent immunosuppressive comorbidities such as solid organ cancer and diabetes mellitus and because the gastric pH seen in patients with advancing age is higher than in younger individuals, making this less acidic environment more tolerable for *L. monocytogenes* (Charlier et al., 2017; Goldstein et al., 2013).

1.5. Virulence factors

Different virulence factors of *L. monocytogenes* have been identified to date. Genes encoding virulence determinants among different strains, serotypes and lineages have been described in the last three decades and in general they are highly conserved, however, some but not all strains harbor virulence-associated polymorphisms and virulence-related sequences that contribute and confer diversity in virulence within the species (Kathariou et al., 2017).

One of the main virulence attributes identified in *L. monocytogenes* is the *Listeria* Pathogenicity Island number 1 (LIPI-1), a 9 kb long region that includes six genes encoding proteins needed in the intracellular cycle. Virulence factors such as listeriolysin O (encoded by *hly*), phospholipases A and B (encoded by *plcA* and *plcB*), zinc metalloproteinase (encoded by *mpl*), actin assembly-inducing protein (encoded by *actA*), and the principal transcriptional regulator of this island, PrfA (encoded by *prfA*) are part of LIPI-1 (Portnoy et al., 1992; Vazquez-Boland et al., 2001; Wernars et al., 1992). Additionally, PrfA also regulates virulence genes not located on LIPI-1, such as the internalin genes *inlA*, *inlB*, and *inlC* important for invasion and cell-to-cell spreading, hence associated with virulence in *L. monocytogenes* (Dramsi et al., 1995; Gaillard et al., 1991). Interestingly, while the gene structure and transcriptional organization of LIPI-1 are identical in the two pathogenic *Listeria* species, *L. monocytogenes* and *L. ivanovii*, there is a significant degree of variance (34-78% identity) when comparing their gene sequences. Likewise, *L. seeligeri*, a nonpathogenic species, carries a nonfunctional LIPI-1 due to an insertion between the *pclA* and *prfA* genes suggesting that LIPI-1 was already present in a recent common ancestor of these species, however, the virulence cluster is totally absent from nonpathogenic species such us *L. innocua*, *L. welshimeri* and *L. grayi*, where its deletion may have occurred at a later stage in the evolution (Bakker et al., 2010; Gouin

et al., 1994; Haas et al., 1992; Vazquez-Boland et al., 2001). Interestingly, natural atypical hemolytic strains of *L. innocua* have been reported, showing that LIPI-1 and *inlA* genes are transcribed, expressed and are functional in the pathogenesis process supporting the existence of a recent common virulent ancestor of *L. monocytogenes* and *L. innocua* (Moura et al., 2019).

Opposite from LIPI-1 that is found in all *L. monocytogenes* strains, other virulence-associated genes have been described for specific strains, serotypes, CCs and lineages (Cotter et al., 2008; Dreyer et al., 2016; Maury et al., 2016a). The *Listeria* Pathogenicity Island number 3 (LIPI-3) for example, is present in certain lineage I strains, particularly in serotypes 1/2b and 4b responsible for human listeriosis outbreaks, and has not been identified in other lineages yet. LIPI-3 groups four genes: *IIsB*, *IIsY,* and *IIsD* that encode enzymes that modify a structural listeriolysin S toxin, a streptolysin S-like hemolysin/bacteriocin encoded by *IIsA* that targets and modulates the host gut microbiota *in vivo* (Kathariou et al., 2017; Quereda et al., 2017b). Additionally, the *Listeria* Pathogenicity Island number 4 (LIPI-4), a cluster of six genes contributing to a putative sugar transport system (cellobiose-family phosphotransferase system-PTS) appears to be unique to CC4 isolates enhancing CNS and MN tropism and accounting for the hypervirulent nature of this CC (Maury et al., 2016a). Other factors also may play a role in the survival of *L. monocytogenes* outside of a host such as the Stress Survival Islet 1 (SSI-1) a cluster of five genes that contributes to the survival in conditions like high salt concentrations and low pH (Hilliard et al., 2018).

Various virulence phenotypes have also been linked to mutations in known virulence genes. More than 30% of isolates in lineage II are virulence-attenuated due to a premature stop codon in *inlA*, while isolates in lineage I harbor a full-length *inlA* (Nightingale et al., 2008). This also is reflected in epidemiological studies where around 96% of the clinical strains express a

full-length functional form of *InlA*, while only 65% of food isolates contained the full-length internalin (Camejo et al., 2011; Maury et al., 2016a). Mutations of *prfA* and *actA* also contribute to attenuated phenotypes observed in *in vitro* assays and are not typically implicated in invasive listeriosis (Rupp et al., 2015). Thus, strain-specific genomic features play key role in virulence and impact aspects like the epidemiology of clinical manifestations, preferred host, and severity of illness and may be associated with specific groups of strains (i.e. sequence types -ST, CC, lineages) as shown for instance by Elinav et al in 2014, where CC2 was found to clearly play a major role in human MN infection cases in Israel (Elinav et al., 2014); by Fox et al, who reported ST204 as the most common ST identified in human clinical infections in Australia (Fox et al., 2016); by Kremer et al, who found an increased incidence of ST6 in CNS infections in the Netherlands (Kremer et al., 2017); or by Dreyer et al, who uncovered ST1 as predominant in ruminant rhombencephalitis cases (Dreyer et al., 2016).

1.6. *Galleria mellonella* as animal model

Listeriosis has traditionally been studied in mice and other species such as gerbils, rabbits, and guinea pigs. Alternatively, model systems such as insects like *Drosophila melanogaster* (Cheng and Portnoy, 2003), nematodes like *Caenorhabditis elegans* (Thomsen et al., 2006), vertebrates like *Danio rerio* (zebrafish) (Levraud et al., 2009) as well as non-human primates (Wolfe et al., 2017) have contributed significantly to study of *L. monocytogenes* infections. However, even when they have given major insights into cell and tissue tropisms, physiology, immune response, and pathophysiology, no optimal animal model of listeriosis has been established and the use of some of these models has remained limited due to poor interaction with cell host receptors, relatively high cost and ethical considerations (Hoelzer et al., 2012; Joyce and Gahan, 2010).

Recent studies have demonstrated that larvae of the insect *Galleria mellonella* are suitable for assessing the pathogenicity and virulence potential of *L. monocytogenes* (Dreyer et al., 2016; Joyce and Gahan, 2010; Kuenne et al., 2013). The main advantage of *G. mellonella* as a biological model is its ability to survive over a range of temperatures (15-37˚C), which is particularly important when studying human pathogens. Furthermore, invertebrate models are cheaper to establish and maintain and are easy to handle facilitating the collection of tissue and/or hemolymph samples due to its relatively large size (12-20 mm) (Cook and McArthur, 2013).

G. mellonella has a more advanced innate immune system compared to other invertebrates. It has both humoral and cellular immunity that share functional homology with the immune system in mammals and can give insights about mammalian infection processes (Joyce and Gahan, 2010; Killiny, 2018). The humoral response uses effector molecules such as melanin, anti-microbial peptides (defensin and lysozyme) and complement-like proteins that immobilize or kill the pathogen. Thus, molecules like the opsonin apolipophorin III with high affinity for bacterial lipopolysaccharides recognize and bind the bacterial cells resulting in an intracellular cascade. Additionally, the cellular response involves phagocytosis, nodulation, and encapsulation and includes the direct intervention of two types of hemocytes that are found circulating in the hemolymph or adhering to internal organs like the digestive tract or the fat body: plasmatocytes, which are large leaf-shaped cells full of lysosomal enzymes, and granular cells with small nucleus and many granules in the cytoplasm (lectins). The phagocytosis process occurs with the recognition of the pathogen by the granular cells followed by the releasing of lectins that promote the attachment of plasmatocytes resulting in the release of lysozymes that break down

the bacterial peptidoglycan layer (Browne et al., 2013; Hoffmann, 1995; Kavanagh and Reeves, 2004).

Although genetic variation within *G. mellonella* populations in addition to variation in environmental and experimental conditions may affect the susceptibility to infection, different studies have produced results that correlate closely with those obtained from other widely used animal models (Cook and McArthur, 2013). Furthermore, both innate immune response and cellular damage can easily be monitored once *G. mellonella* larvae have been infected (Ciesielczuk et al., 2015). Larvae mortality, melanin production, and lactate dehydrogenase (LDH) production have been used as quantifiable indicators of immunogenicity and response along with the evaluation of different time points and changes in bacterial density allowing the assessment of unique virulence patterns of *L.monocytogenes* (Browne et al., 2013; Reddy and Lawrence, 2014).

1.7. Genome structure and genetic diversity of *L. monocytogenes*

One of the first approaches to the genetic diversity of *Listeria* was the publication of the complete genome sequences of *L. monocytogenes* and *L. innocua*. From there, scientists have found that *Listeria* genome is highly conserved and has a strong organization with no inversions or shifts of large regions making its genome very stable. Despite a large number of shared genes, differences exist among *Listeria* genomes, some of which are related to virulence factors and niche adaptation (Buchrieser, 2007; Buchrieser et al., 2003).

In general, the pangenome of a species can be classified in two main groups: core genes that represent the essence and are likely involved with metabolic processes and transcription/translational processes, and accessory genes that represent the variability of the species. *L. monocytogenes* pangenome is estimated to contain ~6,500 genes and it has been

described as “open” meaning that it will continue growing as the number of sequenced genomes increases (Bakker et al., 2010). Genomic analyses have identified ~43% of the *L. monocytogenes* genome as core genes, with the remaining 57% composed of accessory genes (Tan et al., 2015). The accessory genome of different populations within a bacterial species is highly variable due to selective pressures experienced in different environments, however, the introduction of new genetic material from external gene pools seems to be very limited in *L. monocytogenes*. In 2013, den Bakker *et al* studied 21 strains and established that the *L. monocytogenes* genome is compartmentalized with two principal regions: one enriched for core genes and a second region enriched for accessory genes, which was described as a “hot region” for the gain of horizontally acquired information where lineage-specific accessory genes were revealed (Bakker et al., 2013). Likewise, Kuenne *et al* found differences among 16 *L. monocytogenes* strains focused in nine hypervariable hotspots, prophages, three transposons and two mobilizable islands localized specifically in the accessory genome confirming that integration of foreign DNA or generation of genetic diversity by mutation, duplication and recombination is limited (Kuenne et al., 2013).

L. monocytogenes subgroups contain genetic variants like pathogenicity islands, truncated proteins, point mutations, and premature stop codons in virulence or virulence-associated genes that confer increased/decreased virulence to specific strains and might promote different phenotypes and tissue specificity within the same species. Thus, genetic variations can be associated with geographical or niche-specific adaptation, features that have been linked to *L. monocytogenes* epidemiological studies. This is particularly true for specific STs such as ST1, a group of strains associated with neurolisteriosis that have demonstrated to be hypervirulent and hyperinvasive *in vitro* and *in vivo* compared to other genotypes, however even when uncharacterized genetic elements are suspected to be present and contribute to its high virulence,

they have not been identified (Dreyer et al., 2016; Gözel et al., 2019; Rupp et al., 2017). Likewise, ST6 associated with meningitis cases in the Netherlands shows the presence of a phage and an efflux transporter that might confer an increased virulence though decreased susceptibility to disinfectant agents (Kremer et al., 2017); and ST204 a dominant subgroup identify in food and human infections in Australia, shows the presence of mobile genetic elements such as plasmids, phages and transposon insertions indicating a great variation when compare to other genotypes (Fox et al., 2016), hence the significant variation in virulence phenotypes is unlikely to be explained simply by mutations or the presence or absence of the currently known virulence-associated accessory genes (Gözel et al., 2019; Rupp et al., 2017).

While it is currently not possible to predict the virulence of a given isolate based on its genotypic or phenotypic subtype, there is a need to collect more information on strain-specific virulence. Furthermore, the use of strains from diverse niches and associated with specific clinical outcomes or a host is essential to identify virulence factors specific to hypervirulent phenotypes. Many essential genes and genetic elements have been identified in widely used reference strains such as EGD-e, LO28, and 10403S, however putative virulence factors specific to hypervirulent phenotypes might have been overlooked since these reference strains are poorly virulent and invasive (Maury et al., 2016a).

1.8. *L. monocytogenes* whole genome sequencing (WGS)

As WGS became faster and more affordable WGS-based differentiation of *L. monocytogenes* strains has become a useful tool for the investigation of listeriosis around the world. A common application of WGS information is the identification of outbreaks through the comparison of the number and distribution of single nucleotide polymorphisms (SNPs) and the

alignment of SNPs between isolates of interest to determine genomic relationships (Datta and Burall, 2018).

Though the epidemiological approach of WGS has increased our understanding of *L. monocytogenes* population, a whole-genome systematic search of the genetic factors involved in phenotypic properties is still lacking. Genome-wide association studies (GWAS) have been generally conducted from the perspective of the human host to identify and associate genetic variants to specific diseases, thus most of the information and methods for this kind of analysis have been extensively performed in humans(Dutilh et al., 2013). GWAS in bacterial populations though have the potential to improve how we understand, control and treat infectious diseases, therefore, methodological adaptations for the bacterial population have been developed in the last couple of years (Falush, 2016; Falush and Bowden, 2006).

In contrast to human populations, bacterial populations are highly clonal; consequently, when comparing strains from diverse lineages and phenotypes, all variants that separate the lineages may seem to be associated with a phenotype when samples with essentially the same genome are treated as independent, even though there is no causal link with the phenotype evaluated (Power et al., 2017). Additionally, factors such as low rates of recombination may have a confounding effect, which makes the causal SNPs indistinguishable from other linked SNPs, where strong linkage disequilibrium will always restrict the resolution of the approach(Earle et al., 2016). To overcome this, methodological approaches such as the increase of sample size and the use of bioinformatics tools designed to account for clonal populations and low recombination have been developed. treeWAS, for example, is a new phylogenetic method to perform microbial GWAS to distinguish between genetic markers that are truly associated with the phenotype of interest and those that are not with high sensitivity and specificity along

with an easy-to-use interphase compared to other cluster-based techniques (Collins and Didelot, 2018).

1.9. Overall impact

Extensive surveillance programs that include food control and exhaustive investigation of human and animal cases have been implemented due to the high morbidity and case fatality rates associated with listeriosis, however, currently all *L. monocytogenes* isolates are considered as equally virulent by regulatory authorities (S. W. Kim et al., 2018; Maury et al., 2016a). Virulence in *L. monocytogenes* has been demonstrated to be strain dependent, therefore the identification and characterization of genetic elements that help understand prevalence and virulence mechanisms may potentially help to predict the risk in vulnerable populations.

Here, our overall goal was to conduct a genomic and phenotypic assessment of *L. monocytogenes* virulence to collect more information on strain-specific virulence in this species. We first described and compared the genotype distribution of *L. monocytogenes* strains causing disease in ruminants and humans in the Upper Great Plain States and assessed the virulence variation of subgroups in the biological model *Galleria mellonella.* Finally, we investigated the potential association between genetic variants and phenotypes of interest such as clinical outcomes.

In order to apply effective preventive and control measures for *L. monocytogenes*, we need to understand its population structure and identify clinically important genetic variants that might be used as molecular markers for a particular virulence phenotype, since it is likely that strains belonging to the same ST/CC or causing the same clinical manifestation to share unique genetic elements associated with its ability to cause a specific clinical syndrome. Likewise, the use of WGS provides the tools to assess genetic diversity at a high resolution and is increasingly

being used around the world to detect outbreaks and monitor the occurrence of specific clones in human and animal populations, contributing to a more reliable assessment of *L. monocytogenes* epidemiology and a better design of public health programs.

2. GENETIC CHARACTERIZATION OF LISTERIA MONOCYTOGENES FROM RUMINANT LISTERIOSIS FROM DIFFERENT GEOGRAPHICAL REGIONS IN THE U.S[1]

2.1. Abstract

L. monocytogenes infections are an important disease of ruminants worldwide, causing encephalitis, septicemia, and abortions. Ruminant listeriosis can also pose a food safety risk due to the potential for *L. monocytogenes* to enter the food supply via the farm environment. Data on the genetic diversity of *L. monocytogenes* from ruminant clinical cases in the United States is limited. Our goal was to assess the genetic diversity of clinical listeriosis isolates from ruminants in the Upper Great Plains states, a population not well studied, and compares this population to isolates from ruminants in New York State. Multi-locus sequence typing (MLST) was used to classify and compare the genetic diversity of the isolates from the two regions. Loci sequences were compared to all known sequence types using the Pasteur Institute *L. monocytogenes* MLST database. Four novel sequence types (ST) were identified among the Upper Great Plains isolates, and four new STs were classified in the New York collection. Four STs were found to be common across the 2 geographical regions: ST 1, 7, 191, and 204. Strains of ST 7 were most frequently isolated (7/46 isolates). Strains of ST 91 were all associated with fetal infections from the Upper Great Plains. Our results demonstrate that while there are some subtypes commonly found between the two geographic regions, there are also subtypes distinct to each region.

[1] The material in this chapter was co-authored by Abbey J. Steckler, Maria X. Cardenas-Alvarez, Megan K Townsend Ramsett, Neil Dyer and Teresa M. Bergholz. Abbey J. Steckler had primary responsibility for collecting samples, running the MLST analysis and drafted the document. Maria X. Cardenas-Alvarez had primary responsibility for PCR based serotyping, ANI analysis and drafted, and revised all the version of this chapter. Megan K Townsend Ramsett, collected samples and helps with lab methods. Neil Dyer provided the samples and revised the final version of this chapter. Teresa M. Bergholz served revised all the version of this chapter.

2.2. Introduction

Listeriosis is of major veterinary importance primarily in three farm ruminant species: cattle, sheep, and goats. Significant economic losses in livestock production occur due to the high morbidity and mortality in infected animals (Oevermann et al., 2010). Listeriosis manifests as either encephalitis or bacteremia in ruminants, and the septicemic form can lead to fetal infection and subsequent abortions. *L. monocytogenes* is acquired via oral transmission, and outbreaks of listeriosis in herds and flocks are often linked to consumption of contaminated silage (García et al., 2016; Vazquez-Boland et al., 1992; Wiedmann et al., 1994). Additionally, asymptomatic carriage of *L. monocytogenes* in the gastrointestinal tract of ruminants allows the pathogen to multiply and then continue to circulate in the environment (Nightingale et al., 2004).

Up to 50% of fecal samples collected from ruminants without clinical symptoms of listeriosis may contain *L. monocytogenes*, highlighting the potential for spread of the pathogen (Wesley, 1999) . Others that have assessed the prevalence of *L. monocytogenes* on farms have detected this pathogen in 8–22% of water trough or tank samples, 11% of bedding samples, and 8 to 37% of farmyard soil samples, indicating the pathogen can be found throughout the farm environment (Dreyer et al., 2015; Garcia et al., 1996). *L. monocytogenes* is also a human pathogen, causing symptoms similar to those seen in ruminants. Listeriosis in humans is also acquired through the oral route, with consumption of food contaminated with *L. monocytogenes* accounting for 99% of human cases (Scallan et al., 2011). The link between ruminant listeriosis and human listeriosis is not well understood, especially as direct transmission between ruminants and humans rarely occurs. However, ruminants may be an important natural reservoir for *L. monocytogenes* causing human infections, and the pathogen may then enter the human food supply via contamination of foods by manure or water (Oevermann et al., 2010).

Researchers have utilized sequence-based subtyping to characterize *L. monocytogenes* isolates from animal clinical cases, farms, foods, and human clinical cases (Bergholz et al., 2016; Chenal-Francisque et al., 2011; Haase et al., 2014). In an ovine listeriosis outbreak investigation, the same subtype was isolated from infected sheep as well as soil and water samples from the farm, which were considered as potential sources for the outbreak (Dreyer et al., 2015). In an examination of listeriosis isolates from cattle, sheep, and goats in Italy, researchers found that 12/20 isolates were of the same subtype, indicating circulation of that type in the region over time (P. R. D. Rocha et al., 2013). Subtyping of *L. monocytogenes* isolates from dairy herds and farms over time demonstrated that a few specific subtypes of the pathogen persisted over a 6-year period among herd members (Haley et al., 2015). The majority of information on *L. monocytogenes* subtypes in ruminants in the U.S. is from non-clinical isolates (Haley et al., 2015), or based on ribotyping, which is less commonly used than PFGE or MLST (Nightingale et al., 2004; M. A. Pohl et al., 2006). Our goal was to utilize multi-locus sequence typing (MLST) to characterize and compare the isolates causing listeriosis in ruminants from two geographic areas in the U.S., the Upper Great Plains (North Dakota, South Dakota, Minnesota) and New York State.

2.3. Methods

2.3.1. Isolates and DNA extraction

A total of *46 L. monocytogenes* isolates were examined in this study, 19 from New York that were obtained from the Food Safety Lab at Cornell University, Ithaca, NY and have previously been subtyped by ribotyping (M. A. Pohl et al., 2006), and 27 from the Upper Great Plains that were provided by the North Dakota State Veterinary Diagnostic Lab (NDVDL), Fargo, ND. All isolates came from diagnosed clinical cases of listeriosis in ruminants (Table 1).

Isolates were stored at −80 °C in brain-heart infusion (BHI) broth with 15% glycerol, and grown in BHI broth for 20 h prior to use for DNA extraction. DNA was extracted using either the Qiagen DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA) or a modified phenol-chloroform protocol (Flamm et al., 1984). DNA quantity and quality was measured with a Nanodrop (Agilent, Santa Clara, CA). DNA with A230/A260<1.8 were precipitated with 1/10 (vol/vol) 5M sodium acetate and 2.5 (vol/vol) 100% ethanol, followed by suspension of the DNA in 10mM Tris-EDTA. Only DNA with A230/ A260>1.8 was used for PCR.

Table 1. Isolate Information genetic characterization from ruminant listeriosis

Isolate	Source	Clinical	State	Year	Serotype	Lineage	CC	ST
NDS TB-0359	Bovine	CNS	ND	2015	4b (4ab, 4e)	1	1	1
NDS TB-0509	Ovine	CNS	ND	2017	4b (4ab, 4e)	1	1	1
FSL J2-0026	Bovine	CNS	NY	1993	4b	1	1	1
FSL J2-0022	Bovine	CNS	NY	1992	4b	1	1	1239
FSL F3-0695	Bovine	Fetal infection	NY	2004	4b	1	4	4
FSL E1-0054	Bovine	Bacteremia	NY	1999	4b	1	6	6
FSL J2-0036	Bovine	Bacteremia	NY	1986	1/2b	1	59	59
NDS TB-0404	Ovine	CNS	ND	2015	1/2b (3b)	1	191	191
NDS TB-0405	Ovine	CNS	ND	2015	1/2b (3b)	1	191	191
FSL J2-0038	Bovine	CNS	NY	1992	1/2b	1	191	191
FSL R9-0953	Bovine	Bacteremia	NY	2013	4b	1	217	1240
FSL F2-0386	Bovine	CNS	NY	2000	1/2b	1	379	379
NDS TB-0364	Bovine	CNS	MN	2015	4b (4ab, 4e)	1	554	554
FSL E1-0039	Bovine	Fetal infection	NY	2000	1/2b	1	-	1282
NDS TB-0358	Bovine	CNS	ND	2015	1/2a (3a)	2	7	7
NDS TB-0360	Bovine	CNS	ND	2015	1/2a (3a)	2	7	7

Table 1. Isolate Information genetic characterization from ruminant listeriosis (continued)

Isolate	Source	Clinical	State	Year	Serotype	Lineage	CC	ST
NDS TB-0452	Bovine	Bacteremia	MN	2016	1/2a (3a)	2	7	7
NDS TB-0486	Caprine	CNS	ND	2016	1/2a (3a)	2	7	7
FSL E1-0042	Bovine	Fetal infection	NY	2000	1/2a	2	7	7
FSL J2-0007	Bovine	Bacteremia	NY	1989	1/2a	2	7	7
FSL J2-0019	Bovine	Fetal infection	NY	1995	1/2a	2	7	7
NDS TB-0353	Bovine	Fetal infection	SD	2015	1/2a (3a)	2	14	91
NDS TB-0354	Bovine	Fetal infection	SD	2015	1/2a (3a)	2	14	91
NDS TB-0408	Bovine	Fetal infection	ND	2016	1/2a (3a)	2	14	91
NDS TB-0407	Bovine	Fetal infection	ND	2016	1/2a (3a)	2	14	91
NDS TB-0527	Bovine	Fetal infection	ND	2017	1/2a (3a)	2	14	91
NDS TB-0363	Ovine	CNS	MN	2015	1/2a (3a)	2	21	21
FSL J2-0016	Bovine	CNS	NY	1994	1/2a	2	37	37
FSL J2-0002	Bovine	CNS	NY	1992	1/2a	2	89	391
NDS TB-0511	Bovine	Fetal infection	ND	2017	1/2a (3a)	2	121	121
NDS TB-0512	Bovine	CNS	ND	2017	1/2a (3a)	2	121	121
FSL E1-0003	Bovine	CNS	NY	2005	1/2c	2	121	1217
NDS TB-0361	Bovine	Bacteremia	ND	2015	1/2a (3a)	2	199	230
NDS TB-0362	Bovine	CNS	ND	2015	1/2a (3a)	2	199	230
NDS TB-0485	Caprine	CNS	WY	2016	1/2a (3a)	2	204	204
FSL J2-0018	Bovine	Bacteremia	NY	1995	1/2a	2	204	204
FSL J2-0011	Bovine	Fetal infection	NY	1991	1/2a	2	204	204
FSL F3-0526	Bovine	Fetal infection	NY	2003	1/2c	2	412	412
NDS TB-0356	Bovine	CNS	ND	2015	1/2a (3a)	2	451	451
NDS TB-0481	Ovine	CNS	ND	2016	1/2a (3a)	2	451	451

Table 1. Isolate Information genetic characterization from ruminant listeriosis (continued)

Isolate	Source	Clinical	State	Year	Serotype	Lineage	CC	ST
NDS TB-0571	Bovine	CNS	ND	2017	1/2a (3a)	2	451	451
FSL J2-0057	Bovine	CNS	NY	1992	1/2a (3a)	2	918	918
NDS TB-0451	Bovine	CNS	ND	2016	1/2a (3a)	2	-	1057
NDS TB-0453	Bovine	CNS	SD	2016	1/2a (3a)	un	-	1058
NDS TB-0357	Bovine	CNS	ND	2015	4c	un	-	897
NDS TB-0508	Ovine	Fetal infection	ND	2017	4c	un	-	1283

[a] Isolates with 'NDS' were those isolated by the North Dakota State Veterinary Diagnostic Laboratory. Isolates with 'FSL' were obtained from the Food Safety Lab at Cornell University.
[b] Serogroups for the NDS isolates were determined using PCR (Kerouanton et al., 2010), and the corresponding serotypes are reported here. Serotypes for the FSL isolates were obtained from previously reported results (Pohl et al., 2006).
[c] 'un' denotes isolates where the lineage was unable to be assigned.

2.3.2. Serotyping PCR

PCR-based serotyping was performed according to the protocol described by Kérouanton et al. (Kérouanton et al., 2010). This method consists of two PCR reactions, the first a multiplex with six primer pairs, five of which target specific genes for *L. monocytogenes* (lmo0737, lmo1118, *orf*2819, *orf*2110, *prfA*), and one primer pair specific for Listeria spp. (*prs*). A second PCR reaction was performed to target *flaA*, which encodes a flagellar protein present in *L. monocytogenes*. Amplified PCR fragments were separated by 2% agarose gel electrophoresis. Band patterns from the first multiplex PCR were used to classify the isolates into one of five serogroups, and the secondary *flaA* PCR was used to distinguish between serogroups IIa (1/2a, 3a) and Iic (1/2c, 3c).

2.3.3. Multilocus sequence typing (MLST)

MLST based on seven genes was used to classify the multi-locus genotype of each *L. monocytogenes* isolate. Amplification of the 7 loci (*abcZ*, *bglA*, *cat*, *dapE*, *dat*, *ldh*, and *lhkA*)

was conducted using PCR according to Stessl et al. (Stessl et al., 2014). PCR reagents and concentrations were as follows: magnesium chloride (Promega, Madison, WI) at 2.5 mM, dNTPs (Promega) at 200 μM, 5X colorless buffer (Promega) at 1X, GoTaq (Promega) at 1U, and forward and reverse primers (Integrated DNA Technologies, Coralville, IA) at 200 nM each. PCR conditions were: an initial denaturation at 94 °C for 4 min, followed by 35 cycles of 94 °C for 30 s, 52 °C for 30 s, and 72 °C for 2 min. For amplification of *ldh*, the annealing temperature was reduced to 48 °C. Amplified products were visualized on a 1.5% agarose gel, and gel extraction was performed as needed (Omega Gel Extraction Kit, Omega Biotek, Norcross, GA). Amplicons were sequenced using universal primers (Ragon et al., 2008) at Macrogen, USA or McLAB, San Fransisco, USA. Strain information and sequence data are available at the Food Microbe Tracker database, www.foodmicrobetracker.com.

2.3.4. Sequence data analysis

Sequence data were analyzed using Geneious version 6.1 (Biomatters, Auckland, New Zealand). Acquired sequences were aligned and trimmed according to a reference sequence of the appropriate locus obtained from the Pasteur MLST database. A consensus sequence was generated for each locus of each strain, and sequences queried against the Pasteur Institute *Listeria monocytogenes* MLST database. The database was used to assign an allele number to each sequence, and the combination of allele numbers (allelic profile) determines the sequence type (ST). MEGA7 (Kumar et al., 2016) was used to construct a neighbor-joining consensus tree from the concatenated allele sequences, using the Jukes-Cantor model and 1000 replicates. ST 562 from lineage IV was chosen to be the root.

2.3.5. Average nucleotide identity (ANI) analysis

Genomic DNA was extracted as stated above. Quantity of the extracted DNA was assessed using the Nanodrop® Spectrophotometer (Thermo Fisher Scientific, Carlsbad, CA), the Quant-iT™ Picogreen® dsDNA Assay Kit (Thermo Fisher Scientific) and Qbit® fluorimeter (Thermo Fisher Scientific). The Nextera® XT DNA Sample Preparation Kit (Illumina, San Diego, CA) was used for DNA library preparation. Paired-end whole genome sequencing (2×250 bp) was performed on the Illumina MiSeq system at the University of Montana Genomics Center. *De novo* assembly was performed using SPAdes (v 3.10.1) (Bankevich et al., 2012) with the default settings after pre processing the raw reads to remove low quality bases and adapter sequences using Trimmonatic (v 0.32) (Bolger et al., 2014).

A total of 65 reference strains were included in the analysis: 6 strains from different *Listeria* species and 59 *L. monocytogenes* strains from different lineages (28 from lineage 1, 24 from lineage 2, 5 from lineage 3 and 2 from lineage 4) (Table 2). Genome sequences were obtained from GenBank. ANI was calculated using the pyani script v0.2.9 provided at https://github.com/widdowquinn/pyani.

Table 2. *Listeria* spp. strains used for the ANI analysis

Species	Strain ID	Lineage	Accession Number
L. marthii	S4_120	NA	CM001047
L. innocua	Clip11262	NA	NC_003212
L. welshimeri	NCTC11857	NA	NZ_LT906444
L. seeliger	SLCC3954	NA	NC_013891
L. ivanovii- londoniensis	WSLC30151	NA	CP009576
L. ivanovii- ivanovii	NCTC11007	NA	NZ_LT906466

Table 2. *Listeria* spp. strains used for the ANI analysis (continued)

Species	Strain ID	Lineage	Accession Number
L. monocytogenes	CLIP80459	1	NC_012488
L. monocytogenes	F2365	1	NC_002973
L. monocytogenes	FSL J1-220	1	NC_021829
L. monocytogenes	SLCC2540	1	NC_018586
L. monocytogenes	ATCC19117	1	NZ_CP013288
L. monocytogenes	J1816	1	NC_021830
L. monocytogenes	CFSAN023463	1	NZ_CP012021
L. monocytogenes	CLIP 80459	1	NC_012488
L. monocytogenes	L312	1	NC_018642
L. monocytogenes	N2306	1	NZ_CP011004
L. monocytogenes	07PF0776	1	NC_017728
L. monocytogenes	CFSAN006122	1	NZ_CP007600
L. monocytogenes	IZSAM_Lm_hs2008	1	NZ_CP010346
L. monocytogenes	J1926	1	NC_021840
L. monocytogenes	J1817	1	NC_021827
L. monocytogenes	J1776	1	NC_021839
L. monocytogenes	SLCC2378	1	NC_018585
L. monocytogenes	WSLC1042	1	NZ_CP007210
L. monocytogenes	NTSN	1	NZ_CP009897
L. monocytogenes	LL195	1	NC_019556
L. monocytogenes	J2_1091	1	SAMN05326634
L. monocytogenes	SLCC2755	1	NC_018587
L. monocytogenes	R2-502	1	NC_021838
L. monocytogenes	N1-011A	1	NC_021826

Table 2. *Listeria* spp. strains used for the ANI analysis (continued)

Species	Strain ID	Lineage	Accession Number
L. monocytogenes	SLCC2540	1	NC_018586
L. monocytogenes	L2624	1	NZ_CP007686
L. monocytogenes	J2-064	1	NC_021824
L. monocytogenes	CFSAN008100	1	NZ_CP011398
L. monocytogenes	EGD-e	2	NC_003210
L. monocytogenes	10403S	2	NC_017544
L. monocytogenes	08-5923	2	NC_013768
L. monocytogenes	08-5578	2	NC_013766
L. monocytogenes	Finland1998	2	NC_017547
L. monocytogenes	SLCC2479	2	NC_018589
L. monocytogenes	SLCC2372	2	NC_018588
L. monocytogenes	FSL R2- 561	2	NC_017546
L. monocytogenes	Lm60	2	NZ_CP009258
L. monocytogenes	R479a	2	NZ_HG813247
L. monocytogenes	L2074	2	NZ_CP007689
L. monocytogenes	J2-031	2	NC_021837
L. monocytogenes	SLCC7179	2	NC_018593
L. monocytogenes	L2625	2	NZ_CP007687
L. monocytogenes	CFSAN007956	2	NZ_CP011397
L. monocytogenes	C1-387	2	NC_021823
L. monocytogenes	6179	2	NZ_HG813249
L. monocytogenes	J0161	2	NC_017545
L. monocytogenes	SLCC5850	2	NC_018592
L. monocytogenes	EGD	2	NC_022568

Table 2. *Listeria* spp. strains used for the ANI analysis (continued)

Species	Strain ID	Lineage	Accession Number
L. monocytogenes	L2626	2	NZ_CP007684
L. monocytogenes	L2676	2	NZ_CP007685
L. monocytogenes	L1846	2	NZ_CP007688
L. monocytogenes	WSLC1001	2	NZ_CP007160
L. monocytogenes	HCC23	3	NC_011660
L. monocytogenes	SLCC2376	3	NC_018590
L. monocytogenes	M7	3	NC_017537
L. monocytogenes	L99	3	NC_017529
L. monocytogenes	LM850658	3	NZ_CP009242
L. monocytogenes	FSL J1-208	4	NZ_CM001469
L. monocytogenes	PNUSAL002614	4	SRR5016994

2.3.6. Statistical analysis

Fisher's exact test was used to determine significant associations between lineage of an isolate and the clinical manifestation of listeriosis in the ruminant. The number of isolates belonging to each lineage 1 and 2 were allocated to neurologic infection, bacteremia, or fetal infection based on the clinical information provided from the NDVDL or published information on the isolate (Pohl et al., 2006). The three isolates that did not clearly belong to a lineage were excluded from the analysis. Fisher's exact test was implemented in SAS v. 9.4 (SAS Institute, Cary, NC).

2.4. Results

2.4.1. Novel STs identified

MLST was used to classify *L. monocytogenes* isolated from cases of ruminant listeriosis in New York State and the Upper Great Plains (UGP). A total of 26 STs were identified among the 46 isolates (Table 1). Of these 26 STs, 8 (31%) were novel STs that had not been previously identified in the Pasteur Listeria MLST database (Fig. 1).

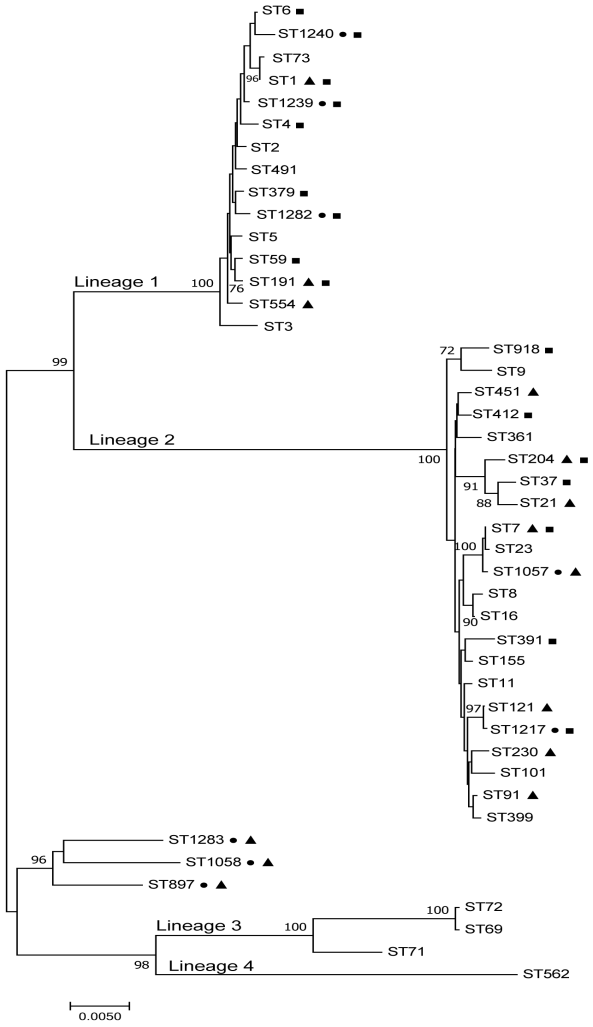

Fig 1. Phylogeny constructed from concatenated MLST loci
Neighbor-Joining method (1,000 bootstraps). Branch tips are labeled with the ST. STs represented in the phylogeny are those identified from this study as well as common STs obtained from the MLST database as references. Novel STs identified in this work are indicated with a circle. STs isolated from ruminant listeriosis cases in New York are indicated with a square; those from the Upper Great Plains are indicated with a triangle.

The 8 novel STs were each represented by a single isolate. Three of the novel STs were single allele variants of known STs, and were assigned to existing clonal complexes (CC). For example, the novel ST 1239 differs from ST 1 at the *abcZ* allele, and is a member of CC 1. Two of the novel STs are singletons, not members of a described CC. These are ST1282 in lineage 1 and ST1057 in lineage 2. Three of the novel STs form a distinct branch between lineage 2 and 3 and do not clearly belong to either lineage (Fig. 1). Notably, these three novel STs all originated from ruminants in North and South Dakota.

2.4.2. Diversity of STs over geographical regions

Overall, lineage 2 isolates (29/46, 63%) were more frequent in our dataset than lineage 1 isolates (14/46, 30%). The isolates from New York State were almost equally represented between lineage 1 (9/19 isolates) and lineage 2 (10/19 isolates), whereas lineage 2 was more frequently isolated from the UGP (19/27 isolates) compared to lineage 1 (5/27 isolates). Serotypes of isolates within each lineage were as expected, with isolates of serotypes ½b and 4b belonging to lineage 1 and isolates of serotypes ½a and ½c belonging to lineage 2 (Table 1).

The majority of the New York State isolates had previously been ribotyped (Pohl et al., 2006) and lineage 2 isolates with the same ribotype tended to have the same ST. For example, isolates FSL E1- 0042, J2-0007, and J2-0019 are all ribotype 1030A, and all belong to ST 7. Isolates FSL J2-0018 and J2-0011 are both ribotype 1039C, and both belong to ST 204. In contrast, lineage 1 isolates of the same ribotype did not have the same ST. For example, isolates FSL J2-0038, F2-0386, and E1-0039 are all ribotype 1042B, but belong to ST 191, ST 379, and ST 1282, respectively. ST 7 was the most common ST in our dataset (7/46 isolates), and was found in isolates from both geographic regions (Table 3). The next most common ST was ST 91, found in 5 isolates from the UGP. Isolates from CC 1, including those of ST 1 and ST 1239,

were found in two cases from each region. ST 451, with 3 isolates, was found only in cases from UGP ruminants. ST 191, ST 204, and CC 121 (ST 121 and 1217) were found in isolates from both New York State and the UGP, with 3 isolates for each of the STs. The remaining STs had only one or two isolates each, and were found only in one of the two geographic regions.

Table 3. Number of isolates by location of isolation and clinical manifestation

Sequence type	Location(s)	Clinical manifestation(s)	Number of isolates
ST 1	ND, NY	Neurologic	3
ST 1239	NY	Neurologic	1
ST 4	NY	Fetal infection	1
ST 6	NY	Bacteremia	1
ST 59	NY	Bacteremia	1
ST 191	ND, NY	Neurologic	3
ST 1240	NY	Bacteremia	1
ST 379	NY	Neurologic	1
ST 554	MN	Neurologic	1
ST 1282	NY	Fetal infection	1
ST 7	MN, ND, NY	Neurologic, bacteremia, fetal infection	7
ST 91	ND, SD	Fetal infection	5
ST 21	MN	Neurologic	1
ST 37	NY	Neurologic	1
ST 391	NY	Neurologic	1
ST 121	ND	Neurologic, fetal infection	2
ST 1217	NY	Neurologic	1
ST 230	ND	Neurologic, bacteremia	2
ST 204	NY, WY	Neurologic, bacteremia, fetal infection	3

Table 3. Number of isolates by location of isolation and clinical manifestation (continued)

Sequence type	Location(s)	Clinical manifestation(s)	Number of isolates
ST 412	NY	Fetal infection	1
ST 451	ND	Neurologic	3
ST 918	NY	Neurologic	1
ST 1057	ND	Neurologic	1
ST 1058	SD	Neurologic	1
ST 897	ND	Neurologic	1
ST 1283	ND	Fetal infection	1

2.4.3. Clinical manifestations by genotype

Each isolate was associated with one of the three different clinical manifestations of listeriosis: neurologic disease, fetal infection, or bacteremia. For lineage 1 isolates, 9 were from neurologic infections, while 2 and 3 were from cases of fetal infection or bacteremia, respectively (Table 1). For lineage 2 isolates, 15 were from neurologic disease, while 10 and 4 were from cases of fetal infection or bacteremia, respectively.

For the three isolates not classified into a lineage, two were from neurologic disease and one from a fetal infection. Significant associations between lineage 1 or 2 and clinical manifestation were identified (Fisher's exact test p value< 0.05). Fetal infections were significantly associated with lineage 2 isolates compared to lineage 1, while neurologic disease or bacteremia cases were not associated with one lineage over the other. The most common CCs from lineage 1 were CC 1 (ST 1 and 1239) and CC 191, and all isolates from these CCs were from neurologic infections. The most common CCs from lineage 2 were CC 7, CC 14 (ST 91),

and CC 204. Isolates from CC 7 and CC 204 were from cases of each of the three clinical manifestations, while all five CC 14 isolates were from fetal infections (Table 1).

2.4.4. Lineage classification of novel STs

An ANI analysis was performed to confirm the lineage classification of three novel STs (ST897, ST1058 and ST1283) from ruminants in North and South Dakota, which formed a distinct branch between lineage 2 and 3 based on the phylogeny constructed from the 7 MLST gene sequences (Fig. 1). By using whole genome sequences we were able to determine the similarity index between the novel STs and the reference genomes (Fig 2). We found that ST1058 belonged to lineage 2 showing identity percentages around 99% when compared to reference genomes from this lineage. Similarly, ST897 and ST1283 were found to belong to lineage 3. When the three novel ST's were compared to different *Listeria* species the average identity percentage was below 95%, as expected for different species.

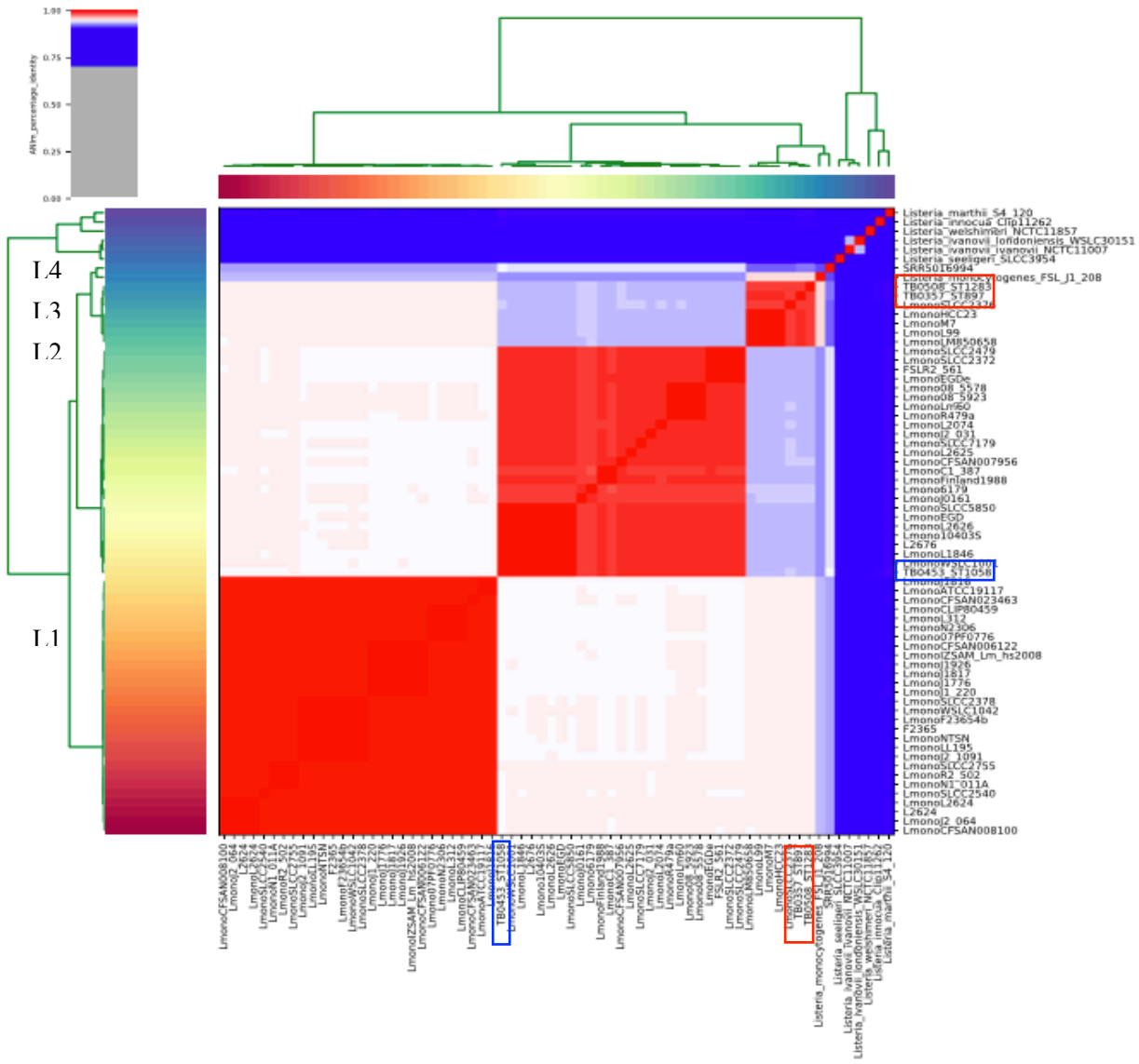

Fig 2. Heat map of ANI analysis to confirm the lineage classification of 3 novel STs
The blue square shows ST1058. The red square shows ST897 and ST1283. An average identity percentage below 95% denotes strains from different species. Average identity percentages closer to 100% are shown in red.

2.5. Discussion

L. monocytogenes lineage 1 and 2 are most frequently isolated from cases of ruminant listeriosis compared to lineages 3 or 4 (Dreyer et al., 2016; Jeffers et al., 2001; M. A. Pohl et al., 2006), which was also observed in our study. Lineage 1 isolates have been more frequently associated with the encephalitic form of listeriosis, where lineage 2 isolates have been more

frequently associated with cases of non-encephalitic infection (Dreyer et al., 2016). In contrast, we found lineage 1 and 2 isolates occurring with similar frequencies from neurologic and bacteremia cases, and lineage 2 isolates significantly associated with fetal infections. Within lineage 1, isolates from CC 1 (also known as epidemic clone 1) have been identified as a significant cause of encephalitis in ruminants in central Europe (Dreyer et al., 2016) and in Italy (P. R. D. Rocha et al., 2013). In a survey of 187 ruminant listeriosis isolates, Dreyer and colleagues found that ST 1, ST 4, and ST 412 accounted for 84% of the listeriosis cases (Dreyer et al., 2016). Here we found that lineage 2 isolates belonging to ST 7 and ST 91 were most frequently isolated from ruminant listeriosis.

Virulence phenotypes have been quantified for some STs of *L. monocytogenes*, with isolates of ST 1, ST 4, and ST 412 found to be hyperinvasive in bovine macrophages compared to isolates of ST 18 and ST 37, which are subtypes more commonly isolated from farm environments (Dreyer et al., 2016). Measuring virulence phenotypes for the frequently isolated STs in this study (ST 7, ST 91) in comparison to the hyperinvasive STs may provide further insights into the virulence capabilities of STs commonly isolated from ruminant listeriosis in the U.S.

The subtype most frequently isolated from ruminant listeriosis, ST 7, has also been isolated from cases of human illness, food, animals and animal feed, and environmental samples, based on a search of ST 7 isolates in the Pasteur MLST *Listeria monocytogenes* database. In contrast, ST 91, the second most frequent subtype that we isolated from ruminant listeriosis, has mainly been isolated from food, animals, animal feed, and the environment, and very rarely from human illnesses. Isolates belonging to CC 1, CC 4, and CC 6 are frequently isolated from cases of human illness, and are considered hypervirulent (Maury et al., 2016a). Isolates of these

subtypes were recovered less frequently in our dataset, indicating they may be not being a significant cause of ruminant listeriosis in the region. Other common STs that we identified, including ST 191 and ST 204, are mainly isolated from food, animals, and the environment, and rarely from human illnesses.

To assign the lineage to three novel isolates that were not classified by MLST, we calculated the average nucleotide identity (ANI). The ANI measure evaluates all orthologous genes shared by a pair of genomes after aligning the sequences and identifying the matching regions. This relatedness index is one of the most robust measurements for microbial taxonomy, replacing the labor-intensive DNA-DNA hybridization (DDH) technique as whole-genome sequencing has become widely accessible (Goodfellow et al., 2014; M. Kim et al., 2014). MLST is a powerful method for genotyping and for establishing phylogenetic relationships among *L. monocytogenes* strains, however, since it uses short fragments from only seven core genes it has limited ability for taxonomic classification. Thus, the use of ANI for whole genome sequences is a useful tool to determine if two genomes belong to the same species and furthermore to the same lineage.

In conclusion, we used MLST-based subtyping to classify and compare *L. monocytogenes* from clinical cases of ruminant listeriosis. Our results indicate that the isolates causing listeriosis in ruminants in the U.S. are genetically diverse, with new sequence types of *L. monocytogenes* still being discovered. Our results also demonstrate that while there are some subtypes commonly found between the two geographic regions, including ST 7, CC 1, ST 121, and ST 204, there are also subtypes distinct to each region. Isolates of ST 91 were the second most common subtype, and were all associated with fetal infections in the UGP. Further research is needed to assess the virulence phenotypes of these frequently isolated subtypes.

3. EVIDENCE OF HYPERVIRULENCE IN LISTERIA MONOCYTOGENES CC14[2]

3.1. Abstract

Listeria monocytogenes is an opportunistic foodborne pathogen that causes central nervous system (CNS) and maternal-neonatal (MN) infections, bacteremia, and gastroenteritis in humans and ruminants. Specific clonal complexes (CC) have been associated with severe listeriosis cases, however, less is known about differences among subgroup virulence patterns.

Galleria mellonella larvae were used to compare virulence phenotypes of 34 *L. monocytogenes* strains representing isolates from CC1, CC6 (from lineage I), and CC7, CC9, CC14, CC37 and CC204 (from lineage II) classified by clinical outcome: BAC, CNS and MN infection. Larvae survival, LD_{50}, cytotoxicity, health index scores and bacterial concentrations post-infection were evaluated as quantifiable indicators of virulence.

Isolates belonging to CC14 and MN-associated infections are hypervirulent in *G. mellonella* as they led to lower *G. mellonella* survival rates and health index scores, as well as reduced cytotoxic effects when compared to other CC and clinical outcomes included here. CC14 isolates also showed increased bacterial concentrations at 8 and 24 h post-infection, indicating ability to survive the initial immune response and proliferate within *G. mellonella* larvae.

Subgroups of *L. monocytogenes* possess different virulence phenotypes that may be associated with niche-specificity. While hypervirulent clones have been identified so far in

---

[2]The material in this chapter was co-authored by Maria X. Cardenas-Alvarez, Megan K Townsend Ramsett, Sahar Malekmohammadi and Teresa M. Bergholz. Maria X. Cardenas-Alvarez had primary responsibility for lab methods and analysis, drafted, and revised all the version of this chapter. Megan K Townsend Ramsett and Sahar Malekmohammadi, help with lab methods. Teresa M. Bergholz revised all the version of this chapter.

lineage I, our data demonstrate that hypervirulent clones are not restricted to lineage I, as CC14 belongs to lineage II. Identification of subgroups with a higher ability to cause disease may facilitate surveillance and management of listeriosis.

Keywords: *Listeria monocytogenes, Galleria mellonella,* virulence, clinical outcomes

3.2. Introduction

Listeria monocytogenes is a facultative intracellular pathogen that causes disease in humans and animals. It is frequently found as a contaminant in food and animal feed (Dhama et al., 2015) and has been associated with different clinical manifestations in humans and farmed ruminants such as central nervous system (CNS) infections, bacteremia (BAC), maternal-neonatal (MN) infections, mastitis and gastroenteritis (Rolhion and Cossart, 2017). To date *L. monocytogenes* isolates represent four phylogenetic lineages that can be subdivided in multiple clonal complexes (CC) and sequence types (ST), which are groups of genetically related isolates that are presumed to have descended from a common ancestor and are assigned using genotypic approaches like multilocus sequence typing (MLST) (Datta et al., 2013; Nyarko and Donnelly, 2015; Wiedmann, 2002).

Within *L. monocytogenes,* variation in virulence among strains has been observed. Epidemiological data indicates that not all strains of *L. monocytogenes* are capable of causing disease, as isolates from only four (out of 13) serotypes identified (1/2a, 1/2c, 1/2b, 4b) are responsible for more than 98% of the human listeriosis reported and are distributed mostly among lineage I and II (Orsi and Wiedmann, 2016). In a national survey in France, strains from CC1, CC2, CC4, and CC6 were identified as the cause of 66% of the maternal-neonatal (MN) infections, 57% of the neurolisteriosis cases and 40% of the septicemia cases (Charlier et al., 2017). Interestingly, the remaining cases were caused by isolates classified as hypovirulent and

intermediate virulent clones through an *in vivo* experiment in mice; these included isolates of CC9 and CC121 from lineage II, associated mainly with food and more often isolated from highly immunocompromised patients (Charlier et al., 2017; Maury et al., 2016a). Although there is evidence of clonal groups characterized by unique virulence or host specificity patterns (Wiedmann et al., 1997), it is not clear yet why some strains are able to cause CNS infections while others are more likely to be associated with either the cause of late-term spontaneous abortions, or bacteremia.

In North America, few studies have described the distribution of *L. monocytogenes* clonal groups. A recent study identified 34 clones detected from a panel of 347 serotype 4b isolates from different sources. CC1, CC6, CC2 and, CC4, all from lineage I, predominated in isolates from human origin, however, in contrast to previous studies conducted in Europe, CC2 was reported as significantly overrepresented among food isolates when compared to human isolates (Lee et al., 2018). Likewise, the epidemiology of *L. monocytogenes* in ruminants varies among regions. In 2017, a study which genetically characterized isolates collected from ruminants in the United States showed that CC7 (ST7) and CC14 (ST91) from lineage II were the most common CC/ST isolated in this host (Steckler et al., 2018), differing from studies conducted in Europe where ST1, ST4 (lineage I) and ST412 (lineage II) were the genotypes most frequently found in ruminants (Dreyer et al., 2016).

Most recently, when serotype 4b isolates from animals were evaluated, 48% belonged to lineage III, followed by CC1, CC2, and CC4 from lineage I (Lee et al., 2018). This suggests not only that virulence potential varies among CC/ST in animal hosts but also that this variation is reflected in the epidemiology of *L. monocytogenes*. This could be an indicator that even when the ST/CCs identified in ruminant infections partially overlapped with those in humans, the

prevalence of CC/ST between human and ruminant are different and might indicate possible differences in pathogenesis, ecology, host adaptation and transmission between CC (Dreyer et al., 2016). A remaining gap in knowledge is whether the association of CC/ST with a higher frequency in listeriosis cases and with a specific clinical outcome could be attributed to higher invasiveness and increased virulence of a group of strains.

In our study, *G. mellonella* was used as a biological model to compare the virulence potential of 34 *L. monocytogenes* strains isolated from human and ruminant hosts as well as non-clinical strains. We determined larvae survival, LD_{50}, lactate dehydrogenase (LDH) production, health index scores, and bacterial concentration post-infection since these features can be used as quantifiable indicators of bacterial pathogenesis and immunogenicity to evaluate clonal groups that may be characterized by unique virulence patterns. Previous studies have demonstrated that the use of invertebrate models is suitable to assess the pathogenicity and virulence potential of bacterial pathogens including *L. monocytogenes* (Joyce and Gahan, 2010; Kuenne et al., 2013; Rakic Martinez et al., 2017; Scalfaro et al., 2017).

3.3. Methods

3.3.1. *G. mellonella* larvae conditions

Larvae were obtained from CritterGrub (Wausau, WI) and Speedy Worm (Alexandria, MN) and stored on wood chips in the dark at 15°C and used within 3 days of receipt. Larvae were selected to be 15–25 mm in length, having a cream color with no grey/black markings or spots.

3.3.2. *L. monocytogenes* strains and culture conditions

This study was conducted with 34 *L. monocytogenes* isolates: five non-clinical isolates and 29 clinical isolates from cases of listeriosis in ruminants (cattle and goat) and humans

collected between 1989 and 2017 by the Food Safety Laboratory at Cornell University, the *Listeria* laboratory at the Center for Disease Control and Prevention (CDC) and the North Dakota State Veterinary Diagnostic Laboratory. We selected representative isolates from CC1, CC6, CC7, CC9, CC14, CC37, and CC204 based on three clinical outcomes (CNS, MN, and BAC) to inoculate *G. mellonella* larvae (Table 1). Bacterial strains were stored in Brain Heart Infusion (BHI) (Hardy Diagnostics, Santa Maria, CA) with 15% glycerol at -80°C. To prepare inoculum, strains were transferred from freezer stocks to BHI agar and incubated for 18 hours at 37°C. One isolated colony was selected and grown in 5 ml BHI broth and incubated under the same conditions. Cells were centrifuged at 2,000x *g* for five minutes and suspended in 1 ml of PBS.

3.3.3. *G. mellonella* killing assay

The suspended culture was serially diluted (up to 10^{-7}) and dilutions were plated in duplicate. We inoculated groups of 10 larvae with 10 μl of six different dilutions of bacteria (10^1 to 10^6 cells/ml), for a total of 60 injected larvae per *L. monocytogenes* isolate. Groups of 10 PBS-inoculated and no-solution-injected larvae were included to control for any lethal effects of the injection process. Infected larvae were incubated for up to 7 days at 37˚C, in the dark, and they were scored every 24 hours for live and dead larvae. Larvae were considered dead when their color turned from pale yellow to dark brown/black and no movement was observed on stimulation (Scalfaro et al., 2017). Three independent biological replicates for each isolate were performed, in separate weeks, using different batches of *G. mellonella* larvae (Ciesielczuk et al., 2015; Kuenne et al., 2013). LD_{50} for each strain was calculated using Probit regression model analysis (Finney and Stevens, 1948).

3.3.4. Lactate dehydrogenase (LDH) assay

The optimal inoculum for the LDH assay was determined by injecting 10 larvae with two different concentrations of *L. monocytogenes* (10^5 and 10^6 cells/ml) where we identified the LD_{50} after 24 hours incubation at 37˚C. Groups of 10 larvae per *L. monocytogenes* strain were inoculated using 10^6 cells/ml as optimal inoculum and then incubated for 4 hours at 37˚C in the dark. After incubation, larvae were chilled on ice for 10 minutes, sterilized using 70% ethanol and sacrificed using a sterile disposable surgical scalpel. Hemolymph was then collected in pre-chilled vials containing crystals of N-phenylthiourea (Sigma-Aldrich, St. Louis, MO) to prevent further melanization. A commercial kit was used to measure LDH (CytoTox 96® Non-Radioactive Cytotoxicity Assay. Promega, Madison, WI). Optical density (OD) was measured at 490nm. PBS-inoculated larvae and larvae injected with no solution were used as negative controls (Ciesielczuk et al., 2015). Likewise, hemolymph from PBS-inoculated larvae was used with the maximum LDH release control (included in the kit) to obtain the maximum amount of LDH that would be released from the cells and calculate the cytotoxicity percentage of each strain as indicated in the commercial kit. Three independent biological replicates for each isolate were performed as described above.

3.3.5. Monitoring *G. mellonella* larvae

Activity, cocoon formation, melanization, and survival were monitored at each time point using the health index scoring system described by Loh *et al* (Loh et al., 2013). A healthy uninfected larva scores between 9 and 10 while a dead larva scores 0.

3.3.6. Enumeration of bacterial cells post-infection

Dilutions were prepared and plated to confirm the initial inoculum. 25 larvae per strain were injected with a target inoculum of 10^6 cells/ml and were incubated at 37˚C in the dark. Four

time points were evaluated: t=0, t=4 hours, t=8 hours, and t=24 hours (t=0 being the initial count right after the inoculation). At each time point 5 larvae per strain were removed from incubation and inspected to determine their health index score (Joyce and Gahan, 2010). Chilled, surface sterilized (70% ethanol), and sacrificed larvae were put into a 50 ml tube containing 10 ml of sterile PBS and 3g of sterile 3mm glass beads and then vortexed for 7.5 minutes. The mixture was serially diluted and plated onto Modified Oxford agar (MOX) (Difco™, Becton Dickinson, Franklin Lakes, NJ) to determine the number of CFU/ml present in each group of larvae at each time point. All experiments were performed in duplicate, in separate weeks, using different batches of *G. mellonella* larvae.

3.3.7. Statistical analysis

We used the Shapiro-Wilk test to verify normal distribution and Bartlett's test for equal variances. We used ANOVA to determine significant differences between CC or outcome for the larval survival, $LD_{50,}$ and cytotoxicity experiments, while the t-test was used to determine significant differences for the health index scores and bacterial counts over time. For multiple comparisons, Tukey's test was used. Correlations between LD_{50} and cytotoxicity percentage was also evaluated by using Pearson correlation. Statistical analyses were performed and graphs were constructed using GraphPad Prism version 8.00 for Mac (GraphPad Software, La Jolla California USA).

3.4. Results

3.4.1. Comparison of larvae survival and LD_{50} values

To compare the virulence potential among strains of *L. monocytogenes* we inoculated *G. mellonella* larvae with six different bacterial concentrations and determined the larvae survival rate, the LD_{50} and the cytotoxicity percentage (LDH) as an indicator of bacterial virulence and

immune response through cell damage in *G. mellonella*. Twenty-nine clinical isolates from CC1, CC6, CC7, CC9, CC14, CC37, and CC204 were selected to inoculate the larvae and represented three clinical outcomes: CNS, MN, and BAC. Five non-clinical isolates from CC7, CC9, and CC14 were also included (Table 4).

Table 4. *L. monocytogenes* strains used for the killing assay and LD_{50}

Clinical Outcome	Clonal complex	Sequence type	Serotype[b]	Strain ID[a]	Source
CNS	1	1	4b	NDS TB0359	Bovine
	1	1	4b	FSL J2-0026	Bovine
	1	1	4b	2014 L-6496	Human
	6	6	4b	2013 L-5275	Human
	6	6	4b	2013 L-5195	Human
	7	7	½a	NDS TB0358*	Bovine
	7	7	½a	NDS TB0360*	Bovine
	7	7	½a	NDS TB0486*	Goat
	37	37	½a	FSL J2-0016	Bovine
	204	204	½a	NDS TB0485	Goat
MN	7	7	½a	FSL E1-0042*	Bovine
	7	7	½a	FSL J2-0019*	Bovine
	14	91	½a	NDS TB0353*	Bovine
	14	91	½a	NDS TB0354*	Bovine
	14	91	½a	NDS TB0407*	Bovine
	14	91	½a	NDS TB0408*	Bovine
	14	91	½a	NDS TB0527*	Bovine
	204	204	½a	FSL J2-0011	Bovine

Table 4. *L. monocytogenes* strains used for the killing assay and LD_{50} (continued)

Clinical Outcome	Clonal complex	Sequence type	Serotype[b]	Strain ID[a]	Source
BAC	1	1	4b	2014 L-6562	Human
	1	1	4b	2014 L-6680	Human
	1	1	4b	2014 L-6708	Human
	6	6	4b	2013 L-5283	Human
	6	6	4b	2015 L-6604	Human
	6	6	4b	2016 L-6178	Human
	7	7	1/2a	FSL J2-0007*	Bovine
	7	7	1/2a	NDS TB0452*	Bovine
	7	7	1/2a	2016 L-6098*	Human
	9	9	1/2c	J5492	Human
	204	204	1/2a	FSL J2-0018	Bovine
Non-clinical	7	85	1/2a	10403S*	Human
	9	210	1/2c	LO28	Faeces
	9	35	1/2a	EGD-e	Rabbit
	9	9	1/2c	2013 L-5605	Food
	14	14	un	FSL H4-0700*	Silage

a, Serotypes as reported in Steckler et al. (Steckler et al., 2018) and Pohl et al. (M. A. Pohl et al., 2006), or by the CDC. 'un' indicates the strain was not serotyped. **b**, Isolate source 'FSL': Food Safety Laboratory Cornell University; 'NDS TB': North Dakota Veterinary Diagnostic Laboratory; Others: Centers for Disease Control and Prevention (CDC). An asterisk following the strain ID denotes strains used for the health index scores and bacterial enumeration post-infection.

When infected with 10^4 cells larvae survival rate varied among CCs over time (up to 7 days) ranging from 29% to 93% (Fig. 3A). Larvae infected with strains from CC9 showed the highest survival rate with a mean of 60% survival seven days post-infection. In contrast, strains from CC14 and CC1 were more virulent and showed lower average survival rates of 29.1% and 33.2%, respectively. CC6, CC7, CC37, and CC204 presented survival rates between 40-50%

after seven days of infection. Additionally, the virulence of *L. monocytogenes* strains was calculated using the CFU/ml derived from plate counts and the number of dead larvae scored every 24 hours, and expressed by the median lethal dose (LD_{50}). Strains from CC14 showed lower LD_{50} values ($\bar{x}$=3.92) compared with strains from the other CC evaluated at all inoculum concentrations, being significantly lower when compared to CC6 ($\bar{x}$=4.94), CC7 ($\bar{x}$=4.71) and CC9 ($\bar{x}$=4.81) resulting in a lower number of bacterial cells needed to cause death in *G. mellonella* (Fig 3B).

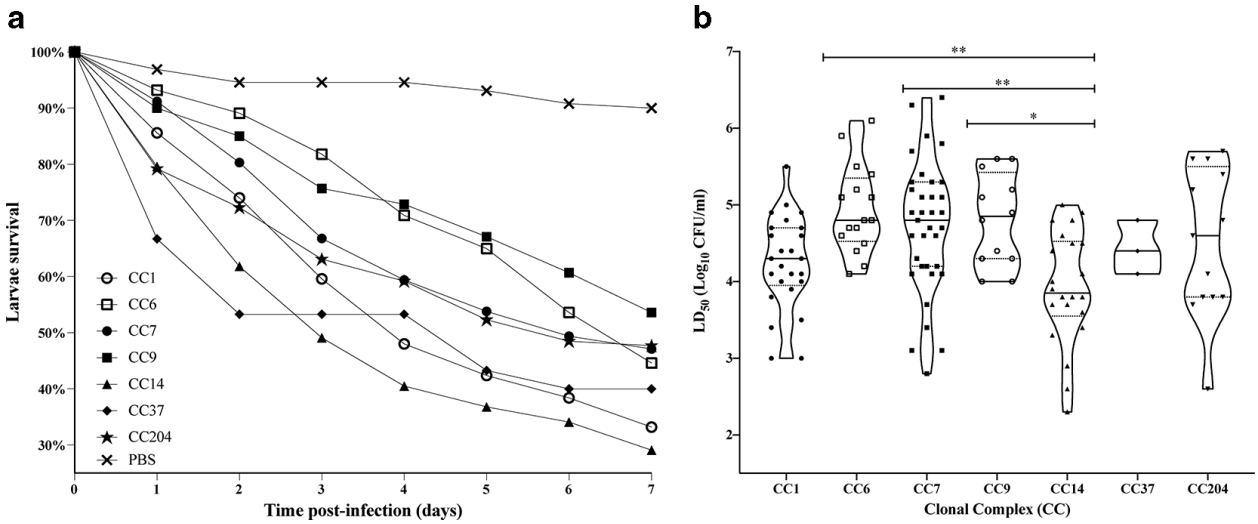

Fig 3. Survival rate and LD_{50} by clonal complex (CC)
Groups of 10 larvae per strain were infected with different doses of *L. monocytogenes* or PBS and were scored every 24 hours for seven days. (a) Survival rate of larvae infected with 10^4 cells/larvae over time. (b) LD_{50} values compared by CC at all inoculum concentrations. Percentage survival and LD_{50} were calculated from 3 separate killing assays 7 days post-inoculation. Bars with asterisks denote statistical significant differences (*=$p<0.05$; **=$p<0.001$).

When grouping the strains by clinical outcome, those isolated from MN infections showed the lowest larval survival rate (32.2%) after seven days, appearing to be more virulent to *G. mellonella* than those isolated from CNS infections, BAC or non-clinical strains (Fig 4A). There was little difference in larval survival rates among isolates from CNS infections (44.1%),

BAC (42.9%), and non-clinical strains (48.7%). In line with the survival rate results, differences in LD_{50} values among clinical outcomes were observed. MN-associated isolates showed LD_{50} significantly lower ($\bar{x}$=4.07) than strains related to CNS infections ($\bar{x}$=4.69) and BAC ($\bar{x}$=4.80), while non-clinical strains showed LD_{50} values in between ($\bar{x}$=4.57), possibly due to factors like the low number of isolates (5) or the CC nature of this group (CC7, CC9, and CC14) (Fig 4B).

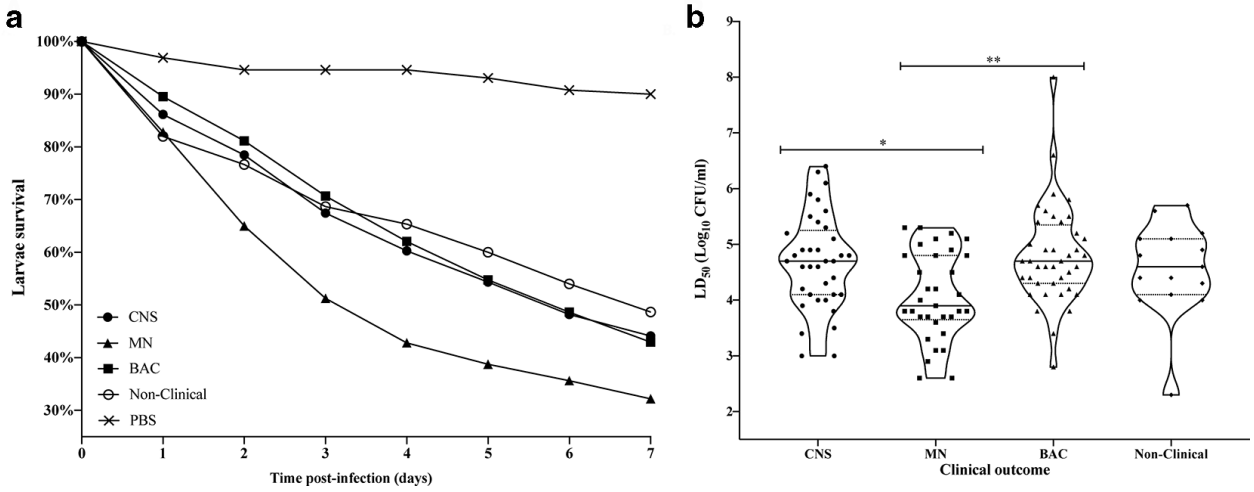

Fig 4. Survival rate and LD_{50} by clinical outcomes
Groups of 10 larvae per strain were infected with different doses of *L. monocytogenes* or PBS and were scored every 24 hours for seven days. (a) Survival rate of larvae infected with 10^4 cells/larvae overtime. (b) LD_{50} values compared by clinical outcomes at all inoculum concentrations. Percentage survival and LD_{50} were calculated from 3 separate killing assays 7 days post-inoculation. Bars with asterisks denote statistical significant differences (*=$p<0.05$; **=$p<0.001$). CNS: Central Nervous System; MN: Maternal-neonatal; BAC: Bacteremia.

3.4.2. Cytotoxicity assessment using *G. mellonella* hemolymph

An increase in the level of LDH after the bacterial invasion is an indicator of host cell damage due to its release from damaged and apoptotic host cells. Larvae were infected with approximately 10^6 cells/larvae, incubated for 4 hours and sacrificed to extract the hemolymph to measure the levels of LDH and calculate cytotoxicity. LDH assay was not conducted on strains from CC6 or CC9 since they showed the highest survival rates. CC14 cytotoxicity was

significantly lower ($\bar{x}$=46.7%) than the observed from CC1 ($\bar{x}$=56.0%), and CC7 ($\bar{x}$=54.7%) (Fig 5A), and when grouped by clinical outcome MN ($\bar{x}$=48.1%) and non-clinical isolates ($\bar{x}$=45.7%) were significantly lower than CNS ($\bar{x}$=56.2%) and BAC ($\bar{x}$=55.5%) (Fig 5B). A significant positive correlation was found when evaluating LD_{50} and cytotoxicity by both CC (R^2= 0.29, P =0.0055) and clinical outcome (R^2= 0.31, P=0.0041), meaning that the increase in one of the variables tends to be associated with an increase in the second variable analyzed (Fig 1. Supplement). These results indicate that strains associated with CC14 and MN infections need fewer bacterial cells to cause disease in *G. mellonella*. In addition, these strains produce less cell damage (low cytotoxicity %) when compared to the other groups evaluated, which may be related to the reduced number of bacterial cells or to an unexplored potential to avoid the host immune response.

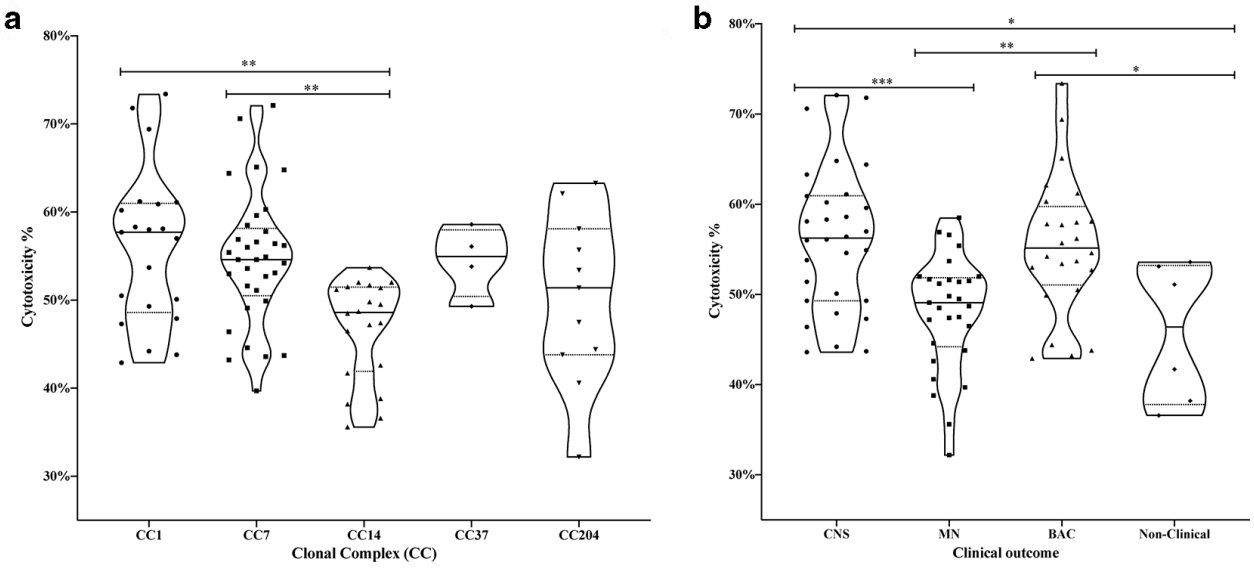

Fig 5. Lactate dehydrogenase **(**LDH) as indicator of host cell damage
Groups of 10 larvae per strain were inoculated using 10^6 cells/larvae as optimal inoculum and then incubated for 4 hours. Cytotoxicity was calculated using the maximum release control to obtain the maximum amount of LDH that would be released from the cells. **(**a) Cytotoxicity comparison by CC. **(**b) Cytotoxicity comparison by clinical outcome. Bars with asterisks denote statistical significant differences (*=p<0.05; **=p<0.001; ***=p<0.0001). CNS: Central Nervous System; MN: Maternal-neonatal; BAC: Bacteremia.

3.4.3. Health index scores (HIS) and bacterial enumeration post-infection

In order to evaluate differences in larvae health and bacterial concentration, health index scores and bacterial cell counts were evaluated at 0, 4, 8 and 24 hours after infection. Higher activity, cocoon formation, and no melanization corresponded to healthier larvae that would score between 9 and 10. In contrast, larvae with minimal to no activity and melanization would score between 2 and 3 (Fig 6).

Fig 6. Health Index Score (HIS)
HIS were assigned according to four observations: activity, cocoon formation, melanization, and survival. (A) HIS of 9-10, (B) HIS of 8, (C) HIS of 6-7, (D) HIS of 5, (E) HIS of 4, (F) HIS of 3, (G) HIS of 0.

A total of fifteen strains from CC7 and CC14 were evaluated based on the previous results, where we found significant differences among these CC that belong to the same lineage II (Table 1). Strains from both CC caused some degree of melanin production relative to the PBS control, but in general strains from CC14 caused a significant decrease of the health index scores at all time points post infection when compared with the strains from CC7, showing scores around 5 after 24 hours (Fig 7A). To determine the growth kinetics of these strains in *G. mellonella*, bacteria were harvested and quantified at each time point. Calculation of the bacterial count difference between the initial counts (t=0) and those at each time point found that *L. monocytogenes* cell numbers decreased for the first 4 hours post-inoculation with both CC7 and CC14, with a rapid increase thereafter. After 8 hours there is a significant difference among the two CC evaluated, where CC14 showed higher counts that correspond with the decreased health

index scores observed before. CC7 failed to recover to the same level as CC14, however, bacterial cells persisted in the larvae after 24 hours (Fig 7B).

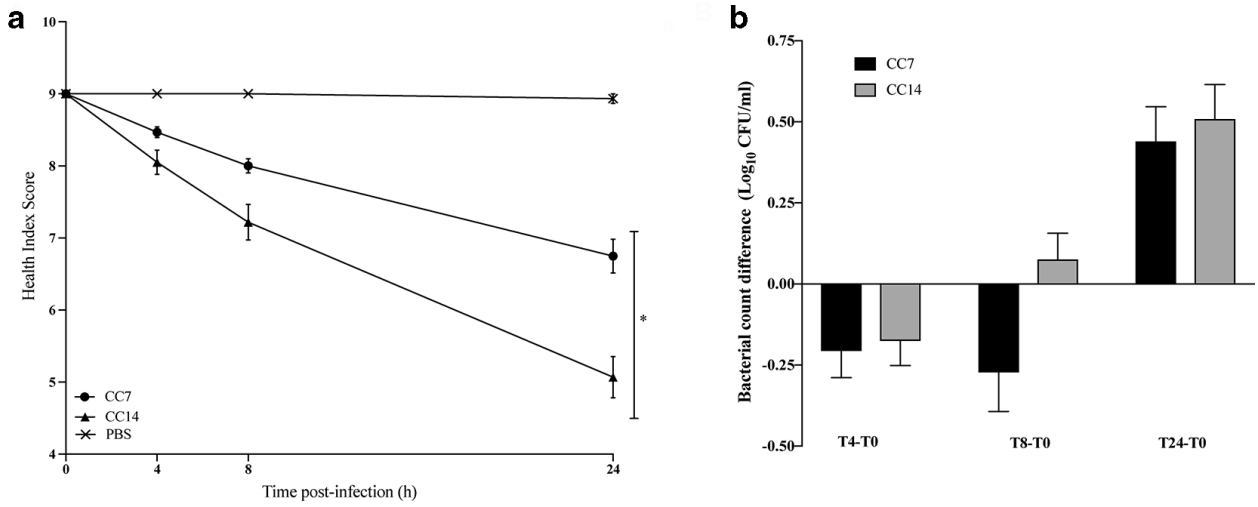

Fig 7. Evaluation of *G. mellonella* larvae inoculated with 10^6 CFU/ larvae at 4, 8, and 24 hours (a) Health index scores based on activity, cocoon formation, melanization, and survival. Each data point represents the mean +/- SEM of larvae post-inoculation. (b) Differences in bacterial numbers between the initial counts (t=0) and subsequent time points. Two separate groups of 5 larvae per strain were tested each time. Bars with asterisks denote statistical significant differences (*=$p<0.05$; **=$p<0.001$).

3.5. Discussion

The use of *G. mellonella* as an insect model to study different human pathogens has been widely used in the last decade (Andrejko et al., 2014; Ciesielczuk et al., 2015; Joyce and Gahan, 2010; Loh et al., 2013; Wand et al., 2013). *G. mellonella* larvae have been demonstrated to be suitable for evaluating bacterial pathogenicity and virulence due to its ability to survive at 37˚C and its advanced immune system, sharing functional characteristics with the immune system in mammals (Killiny, 2018). Here, we assessed larvae survival, LD_{50}, cytotoxicity, health index scores, and changes in bacterial density over time as indicators of humoral and cellular response in the insect larvae after infection with *L. monocytogenes* strains classified by CC and clinical outcomes. *G. mellonella* was found to be a discriminatory model for assessing factors that contribute to virulence differences among *L. monocytogenes* isolates. Furthermore, the results of

assays in the insect larvae agree well with studies in other models of listeriosis like cell cultures, as well as mice, guinea pigs, and gerbils. LD_{50} values in these models are between Log 3 and 8, and LDH ranged from 20-90% depending on the cell type and the animal tested (Golnazarian et al., 1989; Hoelzer et al., 2012; Mukherjee et al., 2010; C. E. Rocha et al., 2017; Roulo et al., 2014; Stelma et al., 1987; Williams et al., 2007).

In the last decade, it has been demonstrated that *L. monocytogenes* subgroups possess unique virulence patterns that can be associated with niche-specific or geographic adaptation impacting the distribution and epidemiology of listeriosis cases (Dreyer et al., 2016; Maury et al., 2016a). This is particularly true for specific subgroups such as ST1, one of the most frequent genotypes globally, associated with neurolisteriosis (Dreyer et al., 2016; Gözel et al., 2019; Rupp et al., 2017), ST6 associated with meningitis cases in the Netherlands (Kremer et al., 2017), and ST204 frequently found in food and human infections in Australia (Fox et al., 2016). Here, we selected strains based on our previous results, where the most commonly isolated ST from ruminant listeriosis cases in the Upper Great Plains states were from ST91 (CC14) and ST7 (CC7), and MN-related isolates were frequent among these two ST (Steckler et al., 2018).

Two mechanisms of *L. monocytogenes* transplacental dissemination have been described to date: the first one is mainly associated with actin-based cell-to-cell spread mediated by the actin-polymerizing protein (ActA) (Bakardjiev et al., 2005; Le Monnier et al., 2007); and the second one is associated with a direct hematogenous invasion through infected phagocytes traveling across the placenta, which involves a major disturbance of the cellular barriers and a strong inflammatory response (Bakardjiev et al., 2005; 2004; 2006; Vázquez-Boland et al., 2017). We found that larvae infected with isolates from CC14 and from MN infections showed an increased virulence with significantly lower larval survival rates, LD_{50}, cytotoxicity, and

health index scores when compared to other CC and clinical outcomes studied here. Furthermore, bacterial counts post-infection at 8 and 24 hours indicated that isolates from CC14 increased significantly after 8 hours, surviving the initial immune response and proliferating within *G. mellonella* larvae. We hypothesize that strains belonging to CC14 and MN- associated may cause low-level infections in *G. mellonella* similar to the first mechanism described, where cell/tissue damage is minimal, allowing successful dissemination while evading the host immune response. This can be explained in our study by the low cytotoxicity levels and increased bacterial counts observed in CC14 and MN groups. Likewise, low LD_{50}, reduced larvae survival and low HIS in these two groups support the hypothesis stated by Vázquez-Boland *et al*. in 2017 that the predominance of one mechanism over the other may depend on the infectious dose and the degree of infection of the host.

Historically, clinical cases were attributed mostly to lineage I strains, while most of the food and environmental strains belong to lineage II, however, recent studies have reported hypervirulent strains associated with invasive listeriosis from lineage II (Charlier et al., 2017; Dreyer et al., 2016; Maury et al., 2016a). Like other strains from lineage II, CC14 has been isolated from diverse sources such as food, animal feed, environment, human and animals. In the Pasteur MLST *Listeria monocytogenes* database, there are 78 isolates from 18 ST belonging to CC14. ST91 is the most represented (37%) followed by ST14 (29%). Similarly, there are 153 cases of MN infections reported, and surprisingly 34% of them are caused by strains that belong to lineage II (https://bigsdb.pasteur.fr/listeria/ last accessed on May 15, 2019). As shown in different epidemiological studies, strains from lineage II such as strains from CC7, CC9, CC37, CC121 and CC412 are included in the group of clones more frequently found as the cause of invasive listeriosis (Aguilar-Bultet et al., 2018; Charlier et al., 2017; Dreyer et al., 2016; Maury

et al., 2016a; Steckler et al., 2018). Even when some of these cases could be explained by the susceptibility of the host (Charlier et al., 2017; Maury et al., 2016a), ST/CC may have specific pathways for infection or tropisms for different tissues or cells. In 2006, Pohl *et al* investigated genetic characteristics and virulence phenotypes of *L. monocytogenes* from 32 cattle in New York. They found that this group of isolates represented two lineages: lineage I predominantly composed of isolates from encephalitis cases, and lineage II composed of an equal number of cases of encephalitis, septicemia, and fetal infection (M. A. Pohl et al., 2006). Furthermore, in 2016, Dreyer *et al* analyzed a total of 187 clinical isolates from ruminants (cattle, goats and sheep) in Europe, describing ST1 (CC1) as predominant, suggesting increased neurotropism of ST1 in these animals and providing evidence of hypervirulent clones causing rhombencephalitis in lineage I and most importantly in lineage II, indicating that this feature is not restricted to strains from lineage I. Little is known about why lineage I strains are usually more virulent than strains from lineage II, yet strains from lineage II have been causing a non-trivial proportion of CNS infections, MN infections, and bacteremia in humans and animals.

In conclusion, after comparing virulence across different *L. monocytogenes* CC, our results indicate that clonal groups are characterized by unique virulence patterns, which might explain the hypervirulence observed in strains from lineage II belonging to CC14, as well as in MN-associated strains when tested in *G. mellonella*. Genetic variants like pathogenicity islands, truncated proteins, point mutations and premature stop codons in virulence or virulence-associated genes have been probed as the source of variation in *L. monocytogenes* subgroups, conferring increased/decreased virulence and promoting different phenotypes and niche specificity within the same species (Rupp et al., 2015). Further studies on genetic variants that

may be linked to the hypervirulent nature of specific CC/ST need to be conducted, as well as those focused on the CC/ST association with specific clinical outcomes.

4. IDENTIFYING GENOMIC FACTORS ASSOCIATED WITH LISTERIA MONOCYTOGENES CLINICAL OUTCOMES

4.1. Abstract

Heterogeneity in virulence potential of *L. monocytogenes* subgroups have been associated with genetic elements that could provide advantages in certain environments to invade, multiply, and survive within a host. Presence of gene mutations has been found to be related to attenuated phenotypes, while presence of groups of genes, such as pathogenicity islands (PI), has been shown association with hypervirulent or stress-resistant clones. To identify genomic elements associated with *L. monocytogenes* strains causing three different clinical outcomes, we evaluated 232 whole genome sequences from invasive listeriosis cases in human and ruminants from the US and Europe. *In silico* MLST was used to classify isolates by clonal complex and sequence type, and a core- single nucleotide polymorphism (SNP) tree using reference strains from the four *L. monocytogenes* lineages was used to classify isolates by lineage, as well as to determine the phylogenetic relationships among the isolates. Sixty-one virulence-associated genes, including four PI, were distributed by CC and unevenly distributed among strains related to bacteremia (BAC), central nervous system (CNS) infections, and maternal-neonatal (MN) infections. Additionally, a total of 10,077 clusters were identified constituting the pan genome of our dataset, and 2,247 present in >95% of the strains were categorized as core genes. Gene-based and SNP-based GWAS were conducted in order to identify locus associated with *L. monocytogenes* clinical outcomes. Orthologous genes of phage phiX174, transfer RNAs and type I restriction-modification (RM) system genes along with SNPs in locus associated with environmental adaptation such as *rpoB* and the phosphotransferase system (PTS) were found associated with one or more clinical outcomes. Detection of phenotype-specific candidate loci

represents an approach that could narrow the group of genetic elements to be evaluated in future studies.

Keywords: Virulence, clinical outcomes, pathogenicity islands (PI), GWAS

4.2. Introduction

Unique virulence features and niche specificity have been described recently for *L. monocytogenes* subgroups. Virulence heterogeneity among *L. monocytogenes* isolates has been reflected as a high frequency of specific clones involved in human and ruminant listeriosis in the US and Europe (Dreyer et al., 2016; Maury et al., 2016a; Rupp et al., 2017). Hypervirulent and hypovirulent clones have been identified along with intrinsic characteristics that lead, for example, to a better survival in the intestinal lumen (Maury et al., 2019), an increased ability to cross host barriers (Aguilar-Bultet et al., 2018; Faralla et al., 2016), or a higher adaptability to food processing environments (Hingston et al., 2017). Some *L. monocytogenes* subgroups are known to cause large outbreaks, while some others are the cause of sporadic cases (Haley et al., 2015).

This diversity in virulence within the species is mainly driven by the presence of groups of genes encoding virulence determinants, as well as polymorphisms among lineages, serogroups and CC (Cotter et al., 2008). *Listeria* Pathogenicity islands (LIPI) such as LIPI-3 present in certain lineage 1 strains (particularly in serotypes 1/2b and 4b), LIPI-4 that appears to be unique in CC4 strains, or the Stress Survival Islet 1 (SSI-1) that contributes to high salt and low pH tolerance, play a role in the survival and enhanced adaptation of specific *L. monocytogenes* subgroups to certain conditions (Hilliard et al., 2018; Kathariou et al., 2017; Maury et al., 2016a; Quereda et al., 2017a). Likewise, mutations in virulence or virulence-associated genes such as

InlA, *prfA,* and *actA* contribute to attenuated phenotypes that impact the listeriosis epidemiology (Camejo et al., 2011; Rupp et al., 2015).

Given the differences of virulence potential among *L. monocytogenes* strains, it is of interest to identify novel genetic variants that might be associated with a particular virulence phenotype, as it is likely that strains causing the same clinical manifestation share unique genetic elements associated with its ability to cause a specific clinical outcome. Genome-wide association studies (GWAS) have been used as a tool to associate genetic variants to specific diseases in humans; however, this method has recently started to be used in bacterial populations (Berthenet et al., 2018; Collins and Didelot, 2018; Fritsch et al., 2019; Maury et al., 2019). GWAS simultaneously assay genetic markers (genes, single nucleotide polymorphisms -SNPs) in the isolates and measures statistical associations between each variant and the phenotype of interest (Brynildsrud et al., 2016; Collins and Didelot, 2018; Q. Wang et al., 2015; Willet and Wade, 2014). This method has been successfully used to identify genomic features associated with host specificity in other bacterial genera such as *Campylobacter* (Sheppard et al., 2013), and more recently to associate *L. monocytogenes* hypervirulent and hypovirulent clones with certain ecological niches (Maury et al., 2019).

Findings in the last few years demonstrate the importance of integrating clinical, epidemiological and experimental approaches to discover new genetic variants associated with clinically important phenotypes. Invasive listeriosis in human and animal populations usually results in CNS infections, MN infections, or BAC, conditions with very high morbidity and case fatality rates (Marder et al., 2018), however, the application of GWAS for the investigation of clinically important phenotypes is still limited. Our goal was to identify genetic markers from a diverse group of listeriosis cases associated with the three principal outcomes caused by *L.*

monocytogenes in order to investigate novel candidate variants that may explain the role of specific genes and SNPs in the pathogenesis processes. Knowing multiple characteristics of strain subgroups such as their CC or ST, in addition to screening for the presence of genetic variants could help food and health agencies to determine why certain isolates might be persisting in a specific host population or more frequently found causing a particular clinical outcome.

4.3. Methods

4.3.1. Strain and genome collection

This study was conducted using a total of 232 *L. monocytogenes* genomes from listeriosis cases associated with three main clinical manifestations (CNS=134, MN=26, and BAC=72) from ruminants and humans. The dataset includes *L. monocytogenes* genomes from in-house collections and the Sequence Read Archive (SRA) database of The National Center for Biotechnology Information (NCBI). The isolates included here represent the phylogenetic diversity present in *L. monocytogenes,* as isolates from the four major lineages were included. Listeriosis cases corresponded to eight different states in the US, as well as from cases in France, Switzerland and Great Britain. Forty-two strains were sequenced specifically for this study, while the remaining 190 were sequenced previously for other studies where the clinical outcome data was available (Dreyer et al., 2016; Maury et al., 2016a) as well as by agencies such as the CDC and the FDA (Table A1). Six *L. monocytogenes* reference genomes belonging to lineages 1: SLCC2540 (NC_018586.1), FSL J1-220 (NC_021829.3); 2: EGD-e (NC_003210.1), 10403S (NC_017544.1); 3: HCC23 (NC_011660.1); and 4: FSL J1-208 NZ_CM001469.1) were downloaded from GenBank (NCBI). Additionally, reference genomes from *Listeria sensu stricto* species such as *L. innocua* CLIP11262 (NC_003212.1), *L. ivanovii* PAM55 (NC_0160011.1), *L.*

seeligeri SLCC3954 (NC_013891.1) and *L. welshimeri* SLCC5334 (NC_008555.1) were also downloaded from GenBank (NCBI) and included when needed.

4.3.2. Phenotype designations

Listeriosis cases were grouped by clinical outcome based on the source of isolation or diagnosis reported. Terms such as rhombencephalitis, encephalitis, meningitis, brain lesion, and brain stem were grouped as Central nervous system (CNS) infection. Maternal-neonatal (MN) infections grouped terms such as abortion, placenta, newborn calf, fetal, and fetus; and terms such as bacteremia, septicemia, blood, liver/lung lesion, and peritoneal fluid were grouped under bacteremia (BAC). Classification by clinical outcome is based on the publicly available data in the databases, as well as data reported to the CDC and FDA.

4.3.3. Whole genome sequencing (WGS)

Isolates were stored at −80 °C in brain-heart infusion (BHI) broth with 15% glycerol, and grown in BHI broth for 20 h prior to use for DNA extraction. Genomic DNA was extracted using either the Qiagen DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA) or a modified phenol-chloroform protocol (Flamm et al., 1984). Quantity of the extracted DNA was assessed using the Quant-iT™ Picogreen® dsDNA Assay Kit (Thermo Fisher Scientific) and the Qbit® fluorimeter (Thermo Fisher Scientific), in addition to the Nanodrop® Spectrophotometer (Thermo Fisher Scientific, Carlsbad, CA). The Nextera® XT DNA Sample Preparation Kit (Illumina, San Diego, CA) was used for DNA library preparation. Paired-end whole genome sequencing (2×250 bp) was performed on the Illumina MiSeq system. Quality control of the reads was performed using FastQC (Wingett and Andrews, 2018) and MultiQC (Ewels et al., 2016) and reads with quality values below *Phred* 20 were excluded from the analysis. *De novo* assembly was performed using SPAdes v. 3.10.1 (Bankevich et al., 2012) with the default settings after pre-processing the raw

reads to remove low-quality bases and adapter sequences using Trimmomatic v. 0.32 (Bolger et al., 2014). The assemblies generated by SPAdes were annotated using PROKKA v. 1.12 (Seemann, 2014) with the default parameters. A BLAST database of annotated *Listeria* strains was generated using 6 reference strains as described in the Prokka manual.

4.3.4. Lineage determination

To classify isolates into genetic lineages, a reference tree based on core single nucleotide polymorphisms (SNPs) was generated using kSNP v. 3.1 (Gardner et al., 2015). Reference genomes for the major lineages of *L.monocytogenes* were included (SLCC2540, FSL J1-220, EGD-e, 10403S, HCC23, FSL J1-208). The resulting maximum parsimony tree (based on the consensus of 100 trees) segregated the four lineages.

4.3.5. *In silico* MLST assignment

To assign isolates to sequence types (ST), *in silico* MLST was performed using the MLST typing tool from the Center for Genomic Epidemiology (https://cge.cbs.dtu.dk/services/MLST/). Clonal complexes (CC) were assigned based on the Pasteur Institute Listeria database (http://www.pasteur.fr/recherche/genopole/PF8/mlst/Lmono.html).

4.3.6. Phylogenetic reconstruction based on core genome SNPs

Variant calling was performed using kSNPs v.3.1 (Gardner et al., 2015) a pipeline that aligns pair-end reads against a reference genome and identifies SNPs in the pan and core genome by a *de novo* assembly, estimating phylogenetic relationships based on them. Parsimony trees are created based on consensus trees from a sample of 100 trees. Kmer size used was k=19 and it was calculated using *Kchooser,* an embedded function in kSNPs. Six reference genomes from the four major lineages were included as mentioned above. The resulted tree was re-rooted using

the reference strain from lineage 4 (FSL J1-208) and edited using iTOL v.4.4.2 (Letunic and Bork, 2016).

4.3.7. Virulence genes screening

A set of sixty-one genes identified as putative or confirmed virulence factors were screened as described in previous studies (Table A2) (Camejo et al., 2011; Maury et al., 2016b). The query gene sequences were extracted from the reference strains EGD-e (NC_003210.1 and AL591981.1), F2365 (NC_002973.6), and CLIP81459 (NC_012488.1). Genomes from reference strains from each *L. monocytogenes* lineage and from other *Listeria sensu stricto* species mentioned above were included along with the 232-listeriosis dataset. CD-HIT-STD-2D (v.4.7) was used to compare the identity between the queries and the listeriosis dataset and to calculate the sequence coverage (parameters -c.90 –n8 –S170) (Li and Godzik, 2006). A gene was considered absent if: 1) the identity percentage between the query and the target sequences was <90%, 2) the gene was not found in the target sequence, 3) the gen was not completed (the maximum difference between the query and the target sequences was established as 170 nucleotides). A gene was considered present if the identity was ≥90%, and the difference between the length of the query and the target sequences was ≤170 nucleotides.

4.3.8. Pan and core genome analyses

Annotated assemblies created by Prokka were taken to calculate the core and accessory genome using Roary (Page et al., 2015). For visualizing and editing the graphs generated we used Phandango v 1.3.0 (http://jameshadfield.github.io/phandango/#/) (Hadfield et al., 2017).

4.3.9. Gene-based GWAS

The accessory gene content of 232 *L. monocytogenes* assemblies was used to perform the gene-based GWAS using treeWAS (Collins and Didelot, 2018), which measures the statistical

association phenotype-genotype while correcting for the confounding effects of clonal population structure and homologous recombination. A file including the phenotypic variables for each individual, the gene presence/absence matrix from the pangenome calculation from Roary (Page et al., 2015), and a phylogenetic tree that accounted for recombination calculated by RAxML v.8.2.10 (Kozlov et al., 2019) (run through CIPRES (Miller et al., 2010) , and ClonalFrameML (Didelot and Wilson, 2015) were used as input data for treeWAS. Locus that were either in <10% or >90% of the strains were not included in this analysis.

4.3.10. SNP-based GWAS

The core SNPs were identified using kSNPs v.3.1 (Gardner et al., 2015) and the output was used to create a matrix of SNPs by position using an *in house* script. SNPs that were either in <10% or >90% of the strains were not included in this analysis. The core SNPs matrix along with the phenotypic variables and the phylogenetic tree calculated by RAxML v.8.2.10 (Kozlov et al., 2019) (run through CIPRES (Miller et al., 2010) and ClonalFrameML were used as input data for treeWAS.

4.4. Results

4.4.1. Variant calling at the core genome level and phylogenetic relationship determination

We identified a total of 18,541 core SNPs by using the completed genome of the reference strain 10403S. A tree based on the SNPs at the core genome level is shown in Figure 8. In the resulting tree, four main branches are observed, corresponding to the different lineages. CC classification showed the 20 most frequently found CC in our dataset in different colors. Most of the strains from the same CC clustered together, however, there are individual strains that are grouped within other CCs, which is the case of some strains from CC1 (LMNC281, LMNC284, LMNC337), CC2 (TB0565), CC4 (LMNC278, LMNC331, LMNC302) and CC6

(LMNC277, LMNC357) from lineage 1. Interestingly, similar to what Aguilar-Bultet *et al* (Aguilar-Bultet et al., 2018) reported, ST14 and ST399 from CC14 (lineage 2) clustered together, however ST91 (cluster formed by a group of 5 isolated from MN infections) was not grouped with the other CC14 strains (ST14/399 included 1 BAC and 5 CNS isolates). These cases may be due to the limited ability of MLST to establish phylogenetic relationships. MLST used fragments from seven core genes unlike the phylogenetic tree showed here, which is based on core genome SNPs; hence discrepancy among the two classification methods may vary when whole genome sequences are used for phylogenetic relationship determination.

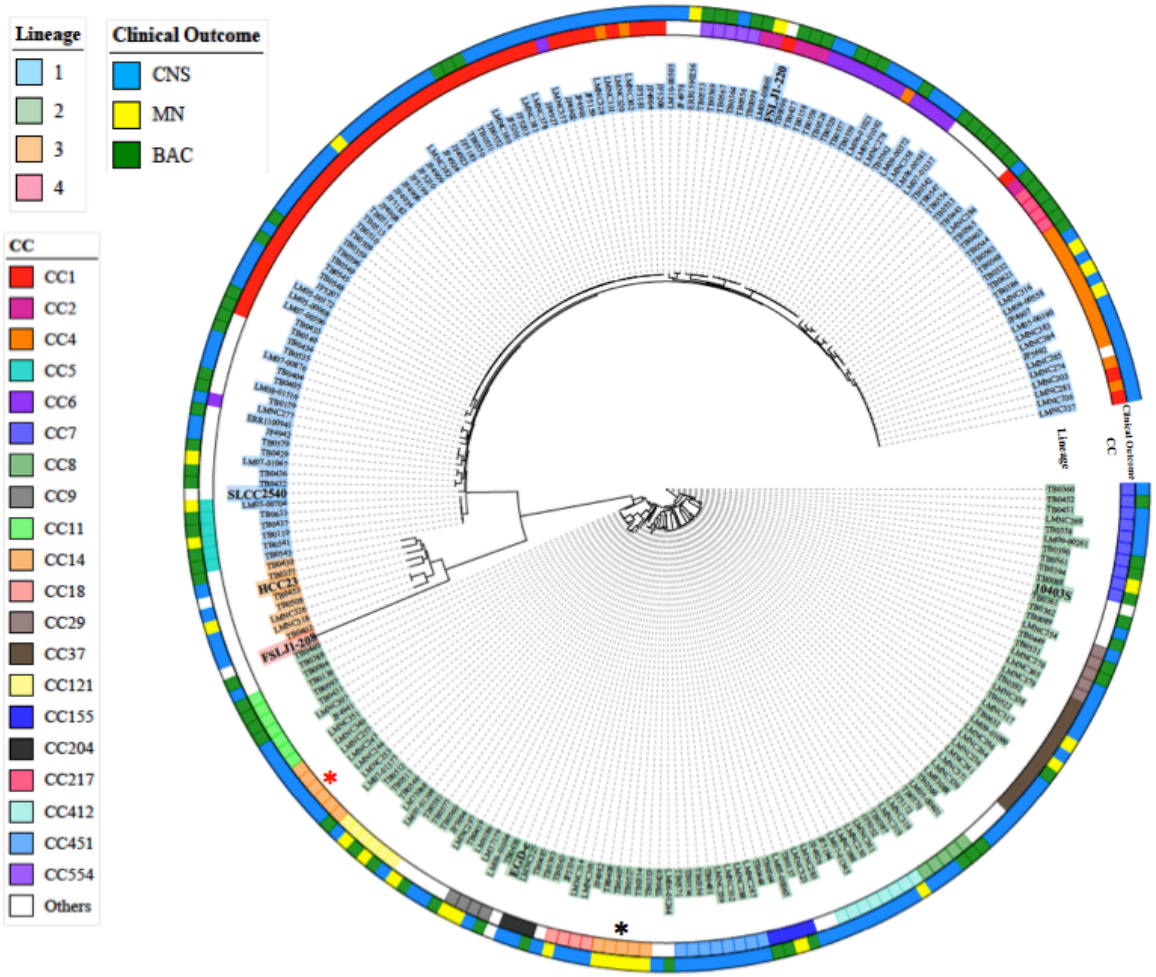

Fig 8. Phylogenetic tree based on the core SNPs

Phylogenetic tree obtained with kSNP3 based on the core SNPs of all 238 strains from lineage 1, 2 and 3, and taking FSL J1-208 from lineage 4 as an outgroup. Colors highlighting the strains' names correspond to different lineages. CC classification is plotted in colors of the inside concentric ring, and clinical outcomes in the outside ring. Red asterisk indicates ST14/399 (CC14), while the black asterisk indicates where ST91 is located.

4.4.2. Distribution of virulence genes from listeriosis cases causing CNS, MN, and BAC

To assess the heterogeneity in virulence among clones, whole genome sequences of 242 strains were analyzed to evaluate the distribution of 61 genetic elements associated with virulence and stress survival (Table A2). Pathogenicity islands LIPI-1, LIPI-3, LIPI-4, and SSI-1 were included in the panel evaluated. Genes encoding LIPI-1 (*actA*, *hly*, *mpl*, *inlA*, *inlB*, *plcA*, *plcB*, and *prfA*) were present in most of the isolates as expected, however, *actA* was found partially present in 45.7% (106/232) of the isolates with gene sequences ~50% shorter than the full gene length reported. 82% (87/106) of the absent/shorter sequences belonged to lineage 1 and 62.3% (66/106) were isolated from CNS infections (Fig 9 and Fig 10). Six genes from LIPI-3 (LMOf2365_1113 to _1118) were also assessed. As described in previous studies, LIPI-3 was absent in lineage 2, 3, and 4, except for one MN-associated isolate from lineage 3 that presented 5 out of the 6 LIPI-3 genes with identity percentages between 97.7% and 99.2% when compared to the reference sequences.

Interestingly, LIPI-4 genes were found to be present not only in CC4 isolates as reported previously, but also in other 14 different CC from lineage 1 including CC1, CC2, CC87, and CC217 showing identities above 99.85%. Furthermore, LIPI-4 was present in lineage 3 (1 human BAC case and 1 ruminant CNS case) and lineage 4 (reference strain J1-208 isolated from a ruminant). SSI-1, a cluster of 5 genes that contributes to the survival of *L. monocytogenes* in stress conditions, was present in a total of 54 isolates (22.3%) mostly from lineage 2 (66.7%).

In general, genes involved in the regulation of virulence such as *prfA*, *sigB*, *VirR*, *hfq*, *SrtA*, and *SecA* were present in all *L. monocytogenes* isolates. Unlike from genes involved in vacuole lysis (i.e. *hly*, *plcA*, *plcB*, *mpl*), intracellular multiplication (i.e. *hpt*, *fri*, *relA*, *OppA*), and evasion and motility (i.e. *sod*, *inlC*, *flaA*), genes involved in adhesion (*ami*, *inlJ*, *inlF*, and *lapB*) and invasion (*aut*, *gtcA*, and *vip*) were more variable among strains: absent, shorter or with identity percentages below 90% in most of the lineage 1,3 and 4 isolates (except for *inlJ* also present in full length in strains from lineage 3), but present in full length in a higher proportion of lineage 2 strains (Fig 9 and Fig 10). This difference among lineages can be explained by the occurrence of alleles sharing low sequence identity as in the case of *inlF*, where the resulting protein share only 74% of the amino acid sequence identity between strains from lineage 1 and 2; or in the case of *inlJ*, where CC1 strains encode for one additional protein domain (Rupp et al., 2017).

CC14 and CC7, which were of special interest as they were found as hypervirulent and hypovirulent respectively, in *G. mellonella* larvae showed differences in virulence gene content when compared. Regarding pathogenicity islands, CC14 harbor only LIPI-1, while CC7 strains included LIPI-1 and SSI-1. Likewise, *Vip*, an invasion-associated gene was absent in CC7 but present in CC14. *inlF* and *inlJ* were present in all the CC7 strains, while only present in strains belonging to ST91 from CC14. Additionally, differences within CC14 sequence types were also apparent, when evaluating *inlF* and *inlJ* present only in ST91, and in i*nlH* exclusively in ST14/399 (Fig 9 and Fig 10). These differences in gene content among STs from the same CC may explain why ST91 and ST14/399 did not cluster together when the core SNP-based phylogeny tree was created (Fig. 8).

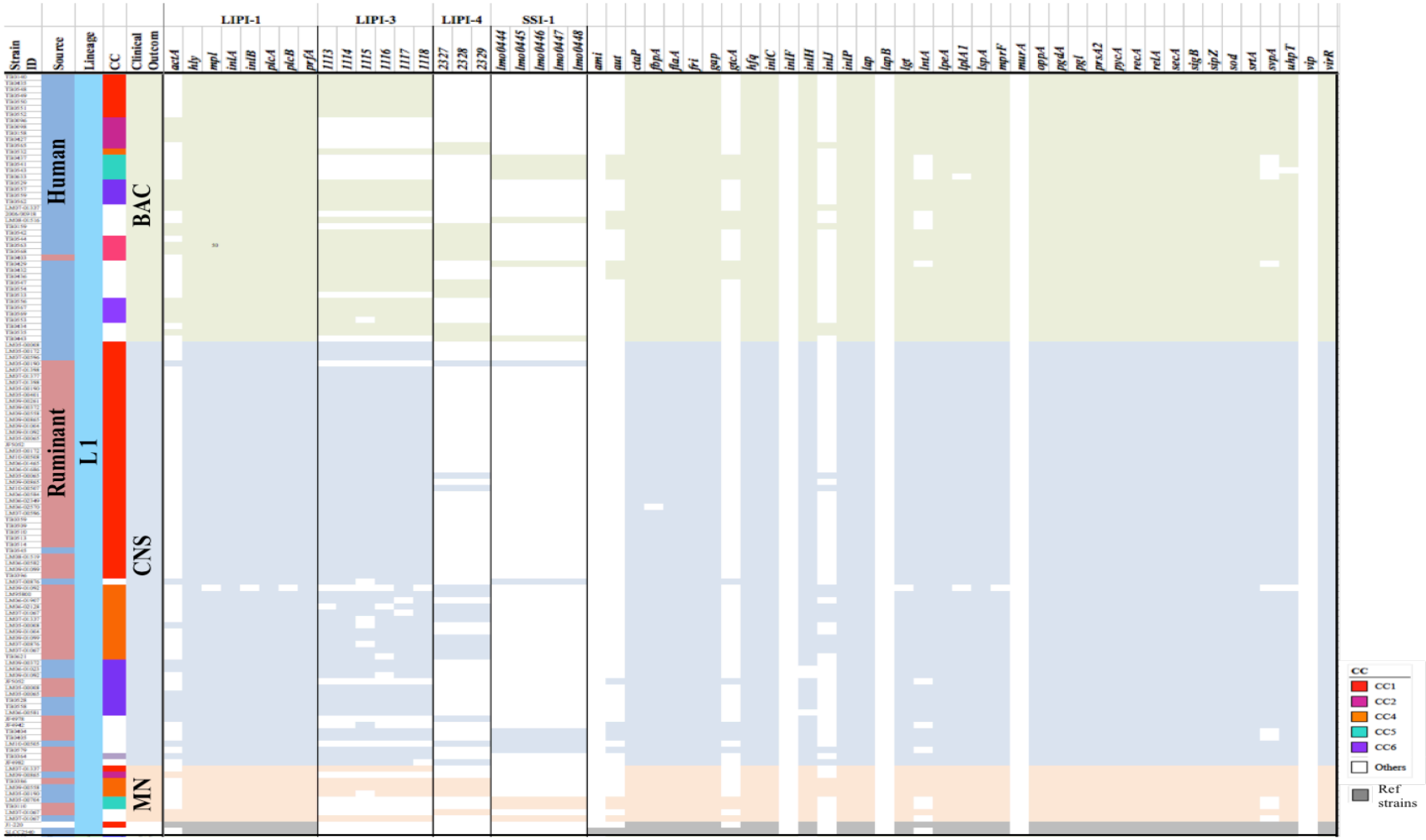

Fig 9. Distribution of 61 genetic elements associated with virulence and stress survival in Lineage 1
Strains grouped by clinical outcomes (BAC, CNS, MN). Colored columns on the right indicate the presence (light green, blue and orange) or absence (white) of the genetic elements

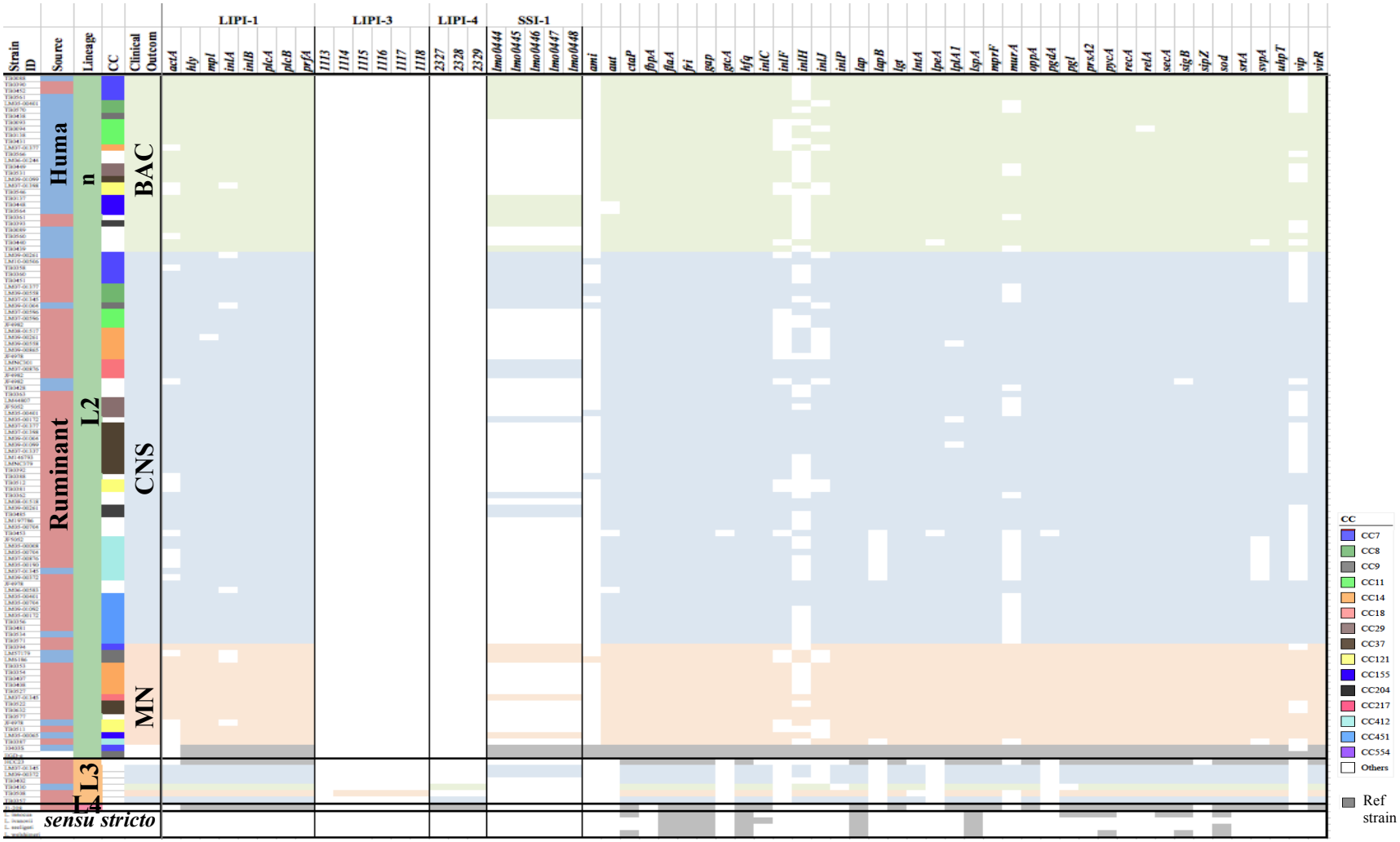

Fig 10. Distribution of 61 genetic elements associated with virulence and stress survival in Lineage 2, 3, and 4
Strains grouped by clinical outcomes (BAC, CNS, MN). Colored columns on the right indicate the presence (light green, blue and orange) or absence (white) of the genetic elements

4.4.3. Pan and core genome analysis

Draft genomes of 232 *L. monocytogenes* strains were used to identify the core and accessory genes in our dataset. A total of 10,077 clusters were found: 1,171 "shell" genes were present in 34-95% of the genomes; 242 "soft-core" genes were present in 95-99% of the genomes; 2,005 core genes were present in 99-100% of the genomes; and 6,659 genes were assigned as "cloud" genes present in less than the 15% of the genomes studied here (Fig 11A). Pan genome size distribution was also calculated according to the number of genomes analyzed: as the number of genomes included in the analysis grew, the pan genome size increased. The pan genome growth can be attributed to the increased number of accessory genes since the number of genes that belong to the core genome remains constant independently of the number of genomes included in the analysis. Although strain-specific genes are identified as accessory genes, the genomes of all *L. monocytogenes* strains seem to be similar in gene content sharing ~2,000 core genes likely involved with metabolic processes, transcription, and translational processes (Fig 11B).

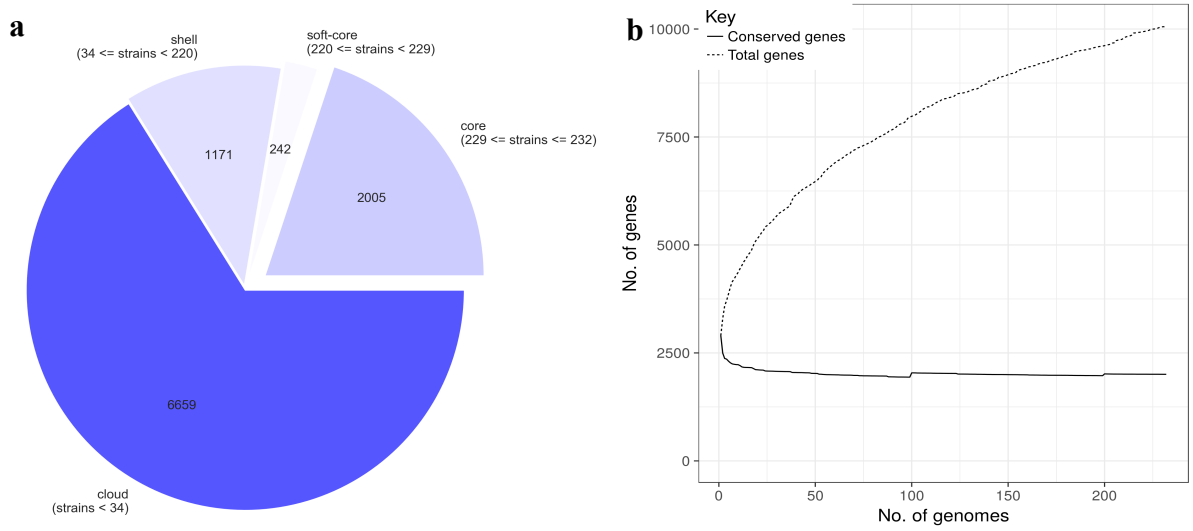

Fig 11. Gene cluster count among 232 *L. monocytogenes* genomes
(a). Classification of gene clusters among the pan genome. (b). Size distribution of pan genome genes related to the number of genomes included in this study.

A phylogenetic tree and a gene presence and absence matrix were also created based on the pan genome calculated by Roary. The core genome represented around 70% (2,005/2,850) of the average number of genes per genome, and close to 20% (2,005/10,077) of the dataset's pan genome (Fig 12). The results indicate extensive clade-specific gene content with genes found only in subsets of strains mainly associated with lineages and CCs.

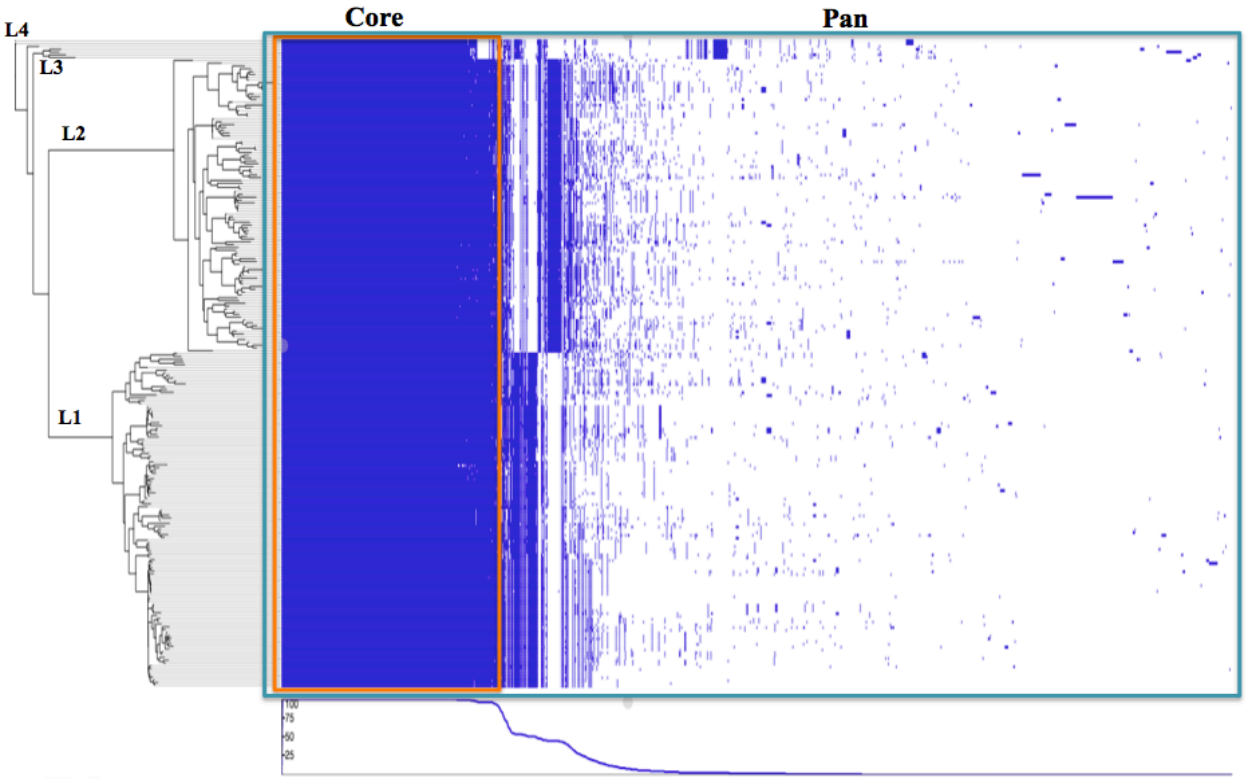

Fig 12. Phylogenetic tree of 232 *L. monocytogenes* isolates compared to a matrix of presence and absence of core and accessory genes
Each row represents a strain's gene content. Each column corresponds to a gene cluster. Columns are ordered by the frequency of gene presence.

4.4.4. Gene association with clinical outcomes

To identify genes associated with CNS, MN and BAC outcomes we performed a genome-wide association study (GWAS) using treeWAS (Collins and Didelot, 2018). 507 genes present among the 10-90% of the isolates were analyzed. Analysis of the accessory gene presence/absence matrix resulted in the identification of a total of 14 genes associated with the

three clinical outcomes evaluated. Ten genes were found to be associated with MN infections, three with BAC and one with CNS infections (this one also found associated with BAC) ($p<0.05$) (Table 5, Fig A2).

Table 5. Identified genes associated with the clinical outcomes evaluated

Locus ID	Gene product	Gene length (ncl)	Clinical Outcome	Number of strains found
EAL09991*	tRNA-Phe(gaa)	74	CNS	195
WP_044683321	hypothetical protein/ phage capsid protein (phiX174)	117	MN	45
ENH11862	hypothetical protein/minor spike protein H (phiX174)	987	MN	45
NP_040712	hypothetical protein/major spike protein G (phiX174)	528	MN	45
ABN49622	hypothetical protein/replication initiation protein gpA (phiX174)	1406	MN	41
WP_016670801	hypothetical protein/phage protein C (phiX174)	144	MN	46
WP_000084700	hypothetical protein/phage protein D (phiX174)	459	MN	45
WP_000033471	hypothetical protein/DNA binding protein J (phiX174)	117	MN	49
EAL09991*	tRNA-Phe(gaa)	73	MN	195
EAL09991	tRNA-Tyr(gta)	83	MN	209
BAO93225	tRNA-Ala(tgc)	75	MN	59
WP_003734550	type I restriction-modification system subunit M	2576	BAC	30
WP_003743526	type I restriction-modification system subunit R	3107	BAC	30
WP_021496534	hypothetical protein	338	BAC	30

* Gene shared by CNS and BAC

BLASTX v. 2.9.0 (Altschul et al., 1997) was conducted using a non-redundant protein database. We established that among MN-associated genes seven of them corresponded to hypothetical proteins that belong to phage phi X 174 (phiX174), a single-stranded DNA virus that infects *Escherichia coli.* Likewise, two genes corresponded to transfer RNAs (tRNA) associated with MN infections (tRNA-Phenilalanina/ Alanina) and one additional associated with MN and CNS infections. These tRNAs may act as integration sites for external genetic elements.

As for BAC associated genes, we identified a gene coding a hypothetical protein of unknown function and two genes related to the type I restriction-modification (RM) system (subunit M and R). RM system recognizes invading foreign DNA and protects bacterial cells against phage infection while reducing horizontal gene transfer, thus it plays an important role in the ecology and evolution of bacteria (Lee et al., 2012).

4.4.5. Core-SNPs association with clinical outcomes

In *L. monocytogenes*, most virulence-associated genes identified belong to the core genome. Therefore, we computed the core genome for our dataset and identify clinical outcome-associated SNPs. A total number of 305,337 SNPs was identified in the pan genome of our *L. monocytogenes* dataset. From there, 18,541 were found in the core genome and 286,796 belonged to the non-core genome. Five clinical outcome-associated SNPs were identified by treeWAS, two of them with CNS cases and the remaining with BAC cases ($p<0.05$) (Table 6, Fig A3).

Interestingly, the SNPs found as associated with the two clinical outcomes are linked to environmental adaptation and virulence of *L. monocytogenes*. One of the SNPs associated with CNS cases (77,529T>C) was found in the gene that codes the phosphoenolpyruvate (PEP) mutase. This enzyme belongs to the isomerases family, specifically the phosphotransferases, whose main function is to transfer phosphate groups and form carbon-phosphorus bonds. Many enzymes that use pyruvate as a substrate are modulated by the phosphotransferase system (PTS) pathway used by bacteria for sugar uptake.

Table 6. Identified SNPs associated with the clinical outcomes evaluated

Associated Position	Reference nucleotide	SNP	Locus tag	Gene product	Protein ID	Region start	Region end	Clinical Outcome
77,529	T	C	LMRG_RS00365	Phosphoenolpyruvate mutase	WP_014600361.1	76760	77533	CNS
271,402	A	G	LMRG_RS01320	DNA-directed RNA polymerase subunit beta	WP_003723046.1	268638	272243	CNS
38,820	C	A	LMRG_RS00170	PTS sugar transporter subunit IIC	WP_003721657.1	38033	39385	BAC
271,402	A	G	LMRG_RS01320	DNA-directed RNA polymerase subunit beta	WP_003723046.1	268638	272243	BAC
298,930	C	T/A	LMRG_RS01435	ABC transporter ATP-binding protein	WP_014600437.1	298775	299440	BAC

Studies have shown that the lack of sigma 54 factor (*rpoN*), a subunit of bacterial RNA polymerase involved in nitrogen and carbon utilization, flagellar synthesis and virulence, directly modifies the PTS pathway affecting the pyruvate to PEP ratio, which influence the expression of pyruvate metabolism-related enzymes (Arous et al., 2004). Additionally, a mutation in one of the genes codifying the PTS sugar transporter subunit IIC (38,820C>A) was found associated with BAC. This system transport carbon sources like glucose and cellobiose, specifically when the bacterium is outside the host cell, which may serve as an environmental signal to switch between an extracellular saprophyte to an intracellular pathogen (S. Wang et al., 2014). The diversity of carbon sources in both hosts and environment and the interaction with other microbes, may be an important factor for *L. monocytogenes* to maintain a diverse repertoire of sugar transporters to cope with these changes.

Likewise, substitutions in *rpoB* that codifies the DNA-dependent RNA polymerase subunit beta (271,402A>G) and in a region of the *ccmA* gene (298,930C>T/A) that encoded an ABC- type multidrug transport system/ATPase component were identified. Mutations in these genes have been linked to rifampicin and rifabutin resistance in strains isolated from food products (Korsak and Krawczyk-Balska, 2016), as well as important for bacterial adhesion, biofilm and lipopolysaccharides biosynthesis (Pieta et al., 2018).

4.5. Discussion

Virulence-associated elements were overlooked for a long time since most of the pathogenesis studies were carried in reference strains where these elements were absent. Reference strains such as EGD-e, 10403S, and LO28 belong to clones occasionally responsible for human clinical cases (CC7 and CC9), hence specific virulence factors associated with hypervirulent clones were commonly missed (Maury et al., 2016a). Only in the last few years,

studies have included a variety of clinical, food and environmental isolates, allowing a better understanding of the *L. monocytogenes* population. In our study, to identify genomic factors associated with *L. monocytogenes* virulence, we collected the genomes of 232 clinical isolates from three main outcomes (BAC, CNS and MN infections) in humans and ruminants from the US and Europe. We integrated phenotypic and genotypic data and identified genetic elements that vary among closely related strains and that may be potentially associated with a *L. monocytogenes* subgroup.

The distribution of 61 virulence-associated genes was surveyed. Genes encoded in LIPI-1were present in most of the isolates, with shorter sequences of *actA* in 45.7% of the isolates, mostly belonging to lineage 1 and associated with CNS cases. ActA is a transmembrane protein that contains two sets of proline-rich- repeats (PRR) and directs three separate events known: 1) actin polymerization independent of repeat regions, 2) initiation of movement dependent on the repeat regions and the amount of ActA, and 3) movement rate also dependent on the PRR (Smith et al., 1996). Deletion of part of the sequence of *actA* gene is a feature that has been related to attenuated phenotypes *in vitro*, as *actA* deletion strains tested in pregnant mice and guinea pigs cause fetal infection with a significant delay, needing a bacterial load 2 log units higher than the wild type virulent strain (Bakardjiev et al., 2005; Le Monnier et al., 2007; Rupp et al., 2017). Here, we found that some strains from i.e. CC1, CC4, and CC6 (lineage 1) described as hypervirulent clones, harbor *actA* mutations that may decrease its ability to spread cell- to- cell causing a reduction in the number of bacterial cells (Rupp et al., 2017).

Likewise, LIPI-3 and LIPI-4 genes were also screened and were not found in isolates from lineage 2, 3 and 4 as reported previously (Hilliard et al., 2018; Maury et al., 2016a; Quereda et al., 2017a), with an exception in an MN-associated isolate belonging to lineage 3,

where 5 of 6 LIPI-3 genes were present. Similarly, LIPI-4 genes associated with CC4, which confers selective tropism for the CNS and fetal-placental organs (Maury et al., 2016a), were found in 14 CC from lineage 1 including CC1, CC2, CC87, CC217, CC382, and CC639, as well as in lineage 3 and 4 strains. This gene cluster was reported initially as exclusive of CC4, however recent studies have reported this pathogenicity island in strains belonging to another lineage 1 CCs (Lee et al., 2018). The role of several genetic elements have been investigated to elucidate the relation of hypervirulence with the increased frequency of certain CC/ST that cause invasive listeriosis, however we must also consider the nature of the host since delayed/hypovirulent clones still can affect patients with highly immunosuppressive comorbidities, as it has been demonstrated before (Charlier et al., 2017; Maury et al., 2016a).

Although, *L. monocytogenes* virulence-associated genes present variations at the genetic level within the species, it is still considered as highly clonal sharing around 70% of the genes among strains. The increased number of accessory genes (around 80% of the pan genome) is responsible for most of the strain-specific features even when variation in core virulence genes is present. Therefore, the accessory genes and the core SNPs were evaluated to identify associations with BAC, CNS, and MN outcomes. We identified orthologous genes of phage phiX174 associated with MN infections. Prophage genes have been found to confer increased virulence in *L. monocytogenes* strains in murine models, also playing an important role in niche adaptation. There are three main bacteriophages that infect *L. monocytogenes*: A006, A500, and P35. These phages have genomes between 35.8 and 134.5 kb in size, with G+C contents between 35.5 and 40.8%. They belong to the Siphoviridae family and feature a similar genome organization, where open reading frames (ORFs) are organized into functional clusters that reflect the direction of transcription. Furthermore, integration sites in the bacterial genome

revealed that both A500 and A006 specifically target the 3' ends of tRNA genes (Hain et al., 2007).

Analysis of genomes from ST204 or ST121, for example, have shown that the majority of variations are linked to mobile elements such as plasmids, transposons and phage insertions and that this elements were conserved in the ST population, suggesting that these elements may provide advantages in the diversity of niches were these STs are found (Fox et al., 2016; Schmitz-Esser et al., 2015). Additionally, prophage regions were located adjacent to tRNAs, indicating that tRNAs are anchoring elements for the uptake of prophage DNA (Hain et al., 2012). Diverse bacterial communities in the 'farm to fork' environments can also influence genetic diversity and contribute to the fact that genetic elements such as phages could be horizontally transferred, conferring new functions like resistance to phagocytosis by macrophages, drug/sanitizer resistance mechanisms, increased biofilm formation or adhesion to human cells (Bergholz et al., 2016; Tinsley et al., 2006).

We also identified genes coding the type I RM system, a group of DNA methyltransferases (MTases) that protect bacterial cells. This system can target foreign invading DNA with restriction endonucleases, and associated methyltransferases protect host DNA from restriction. It is also involved in the regulation of gene expression helping bacteria to cope with environmental changes in nutrient availability, pH, temperature and osmolarity (Pieta et al., 2018). Additionally, SNPs in genes associated with the PTS sugar transporter were identified. This system uptakes carbon sources and is active principally outside the cell, playing an important part in environmental signaling and niche adaptation (S. Wang et al., 2014).

Overall, treeWAS was able to identify genes and SNPs in significant association with three clinical outcomes. Subsequent analyses are required to confirm that a causal relationship is

truly present and verify the role of the associated genes with the pathogenesis process of each clinical outcome. treeWAS was able to control for population structure and recombination while still showing a high statistical power to detect associations in a clonal bacterial population. Similar approaches have been already applied to find associations between genetic variants and antimicrobial resistance (Suzuki et al., 2016), growth in cold conditions (Fritsch et al., 2019) and host adaptation (Maury et al., 2019), therefore a candidate gene approach to screen for genetic markers associated with a phenotype of interest would provide with a narrow set of elements that could be evaluated in the future. Moreover, the detection of isolates harboring phenotype-specific genetic markers could potentially reinforce preventive and control measures.

5. OVERALL CONCLUSIONS

Virulence differences between *L. monocytogenes* subgroups are not only due to single nucleotide differences or the presence or absence of a gene but also due to a different gene composition, gene expression and protein polymorphisms (Aguilar-Bultet et al., 2018; Falush and Bowden, 2006). Environmental conditions such as temperature, osmotic stress, and pH perturb the physiological balance of microorganisms and modulate its response so it can cope with the environmental variation. Thus, phenotypes of genetically identical individuals growing under different conditions can differ phenotypically indicating phenotypic plasticity for particular traits.

An example of phenotypic plasticity in *L. monocytogenes* is observed under salt stress, when rod-shaped cells become elongated at high salt concentrations (12.5%) likely due to the *minC* gene over-expression (Kale et al., 2017). Similarly, genome expression studies during mouse infection demonstrated that ~20% of the genome is differentially expressed through gene activation compared to growth in rich broth medium, coordinated by a complex regulatory network with an important role of virulence regulators PrfA, SigB and VirR (Camejo et al., 2009). Phenotypic plasticity involves a change in morphology, physiology, and/or behavior lead by a change on gene expression. Therefore, this fact expands the set of potential genetic determinants to elucidate *L. monocytogenes* patterns to cause different clinical manifestations.

A key point is that virulence is a process that depends on the interaction of the pathogen with the host environment rather than being an independent event, and where a dynamic cross-talk between the host and the pathogen has to occur (Camejo et al., 2009; Poulin and Combes, 1999). Whole genome analysis allow the identification of unique bacterial genes for certain *L. monocytogenes* subgroups, which may be critical for the infection process reflected as different

clinical outcomes, however whole genome expression studies within the host may explain why some bacterial subgroups are able to cause CNS infections while others are the cause of BAC or MN infections.

6. FUTURE STUDIES

Data presented in this study demonstrate the virulence diversity within *L. monocytogenes* strains. Variation in epidemiology distribution, differences of virulence potential in *G. mellonella* and heterogeneity in genomic patters confirmed that this microorganism possesses a wide repertory of mechanisms to adapt and cause disease. Based on our results, virulence genes of the LIPI-1, LIPI-3, LIPI-4 and SSI-1 islands as well as *internalins* from our dataset need to be analyzed in more detail in order to detect specific mutations and the origin of size variations of genes evaluated here. Resequencing of the strains available by alternative methods that provide longer reads will allow a more accurate assembly and mapping for the evaluation of genetic determinants.

Likewise, identification of resistance genes such as *tetM* and *tetS* for tetracyclin, *penA* for penicillin, and the *bcrABC* locus, the *Tn6188-qac* transposon, the efflux pumps *emrE* and *quaA* for quaternary ammonium sanitizers and antiseptics (such as benzalkonium chloride), as well as mutations in *rpoB* for rifampicin/rifabutin resistance, could be assess *in silico* to compare the differences in proportions of virulence and resistance markers between clinical outcomes, 2) to establish the level of variability specific to each outcome, and 3) to determine if the presence of resistance genes may be involved in a better ability of some clones to cause different outcomes of invasive listeriosis. Additionally, the identification of gene candidates potentially involved in virulence and associated with BAC, CNS, and MN infections could be studied further by reverse genetics. Generation of mutants by transposon sequencing (TnSeq) to assign and validate specific functions as well as to evaluate their impact on virulence in animal models, such as the one standardized here will contribute to a better understanding of the role of these candidate genes.

Lastly, we account for the fact that *L. monocytogenes* variability is not limited to genetic differences among strains. Environmental conditions, physiology, and epigenetics, as well as host immune conditions and bacterial dose ingested (among others), are associated with listeriosis manifestations (Bergholz et al., 2018). *In vivo* "omics" studies such as gene expression profiles will help to elucidate the source of the *L. monocytogenes* variation, including potential differences in genotypes associated with different outcomes of invasive listeriosis and host tropism. The use of cell cultures such as human placental fibroblast and trophoblast progenitor cells, the bovine macrophage line BoMac, as well as fetal bovine brain and epithelial cells may be useful for the study of differences in gene expression of outcome-specific strains. It would be worthy to explore further if strains derived from bacteremia are less virulent than strains capable to cross the blood-brain barrier or those that cross the blood-placental barrier to become implicated in CNS and MN infections, respectively.

REFERENCES

Aguilar-Bultet, L., Nicholson, P., Rychener, L., Dreyer, M., Gözel, B., Origgi, F.C., Oevermann, A., Frey, J., Falquet, L., 2018. Genetic Separation of Listeria monocytogenes Causing Central Nervous System Infections in Animals. Front Cell Infect Microbiol 8, 20. doi:10.3389/fcimb.2018.00020

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402. doi:10.1093/nar/25.17.3389

Andrejko, M., Zdybicka-Barabas, A., Cytryńska, M., 2014. Diverse effects of Galleria mellonella infection with entomopathogenic and clinical strains of Pseudomonas aeruginosa. J. Invertebr. Pathol. 115, 14–25. doi:10.1016/j.jip.2013.10.006

Arous, S., Buchrieser, C., Folio, P., Glaser, P., Namane, A., Hébraud, M., Héchard, Y., 2004. Global analysis of gene expression in an rpoN mutant of Listeria monocytogenes. Microbiology 150, 1581–1590. doi:10.1099/mic.0.26860-0

Bakardjiev, A.I., Stacy, B.A., Fisher, S.J., Portnoy, D.A., 2004. Listeriosis in the pregnant guinea pig: a model of vertical transmission. Infect. Immun. 72, 489–497. doi:10.1128/iai.72.1.489-497.2004

Bakardjiev, A.I., Stacy, B.A., Portnoy, D.A., 2005. Growth of Listeria monocytogenes in the guinea pig placenta and role of cell-to-cell spread in fetal infection. J. Infect. Dis. 191, 1889–1897. doi:10.1086/430090

Bakardjiev, A.I., Theriot, J.A., Portnoy, D.A., 2006. Listeria monocytogenes traffics from maternal organs to the placenta and back. PLoS Pathogens 2, e66. doi:10.1371/journal.ppat.0020066

Bakker, den, H.C., Cummings, C.A., Ferreira, V., Vatta, P., Orsi, R.H., Degoricija, L., Barker, M., Petrauskene, O., Furtado, M.R., Wiedmann, M., 2010. Comparative genomics of the bacterial genus Listeria: Genome evolution is characterized by limited gene acquisition and limited gene loss. BMC Genomics 11, 688. doi:10.1186/1471-2164-11-688

Bakker, den, H.C., Desjardins, C.A., Griggs, A.D., Peters, J.E., Zeng, Q., Young, S.K., Kodira, C.D., Yandava, C., Hepburn, T.A., Haas, B.J., Birren, B.W., Wiedmann, M., 2013. Evolutionary Dynamics of the Accessory Genome of Listeria monocytogenes. PLoS ONE 8, e67511–16. doi:10.1371/journal.pone.0067511

Bakker, den, H.C., Warchocki, S., Wright, E.M., Allred, A.F., Ahlstrom, C., Manuel, C.S., Stasiewicz, M.J., Burrell, A., Roof, S., Strawn, L.K., Fortes, E., Nightingale, K.K., Kephart, D., Wiedmann, M., 2014. Listeria floridensis sp. nov., Listeria aquatica sp. nov., Listeria cornellensis sp. nov., Listeria riparia sp. nov. and Listeria grandensis sp. nov., from agricultural and natural environments. Int. J. Syst. Evol. Microbiol. 64, 1882–1889. doi:10.1099/ijs.0.052720-0

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477. doi:10.1089/cmb.2012.0021

Bergholz, T.M., Bakker, den, H.C., Katz, L.S., Silk, B.J., Jackson, K.A., Kucerova, Z., Joseph, L.A., Turnsek, M., Gladney, L.M., Halpin, J.L., Xavier, K., Gossack, J., Ward, T.J., Frace, M., Tarr, C.L., 2016. Determination of Evolutionary Relationships of Outbreak-Associated Listeria monocytogenes Strains of Serotypes 1/2a and 1/2b by Whole-Genome Sequencing. Applied and Environmental Microbiology 82, 928–938. doi:10.1128/AEM.02440-15

Bergholz, T.M., Shah, M.K., Burall, L.S., Rakic Martinez, M., Datta, A.R., 2018. Genomic and phenotypic diversity of Listeria monocytogenes clonal complexes associated with human listeriosis. Applied Microbiology and Biotechnology 102, 3475–3485. doi:10.1007/s00253-018-8852-5

Berthenet, E., Yahara, K., Thorell, K., Pascoe, B., Meric, G., Mikhail, J.M., Engstrand, L., Enroth, H., Burette, A., Megraud, F., Varon, C., Atherton, J.C., Smith, S., Wilkinson, T.S., Hitchings, M.D., Falush, D., Sheppard, S.K., 2018. A GWAS on Helicobacter pylori strains points to genetic variants associated with gastric cancer risk. BMC Biol. 16, 84. doi:10.1186/s12915-018-0550-3

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. doi:10.1093/bioinformatics/btu170

Borucki, M.K., Reynolds, J., Gay, C.C., McElwain, K.L., Kim, S.H., Knowles, D.P., Hu, J., 2004. Dairy farm reservoir of Listeria monocytogenes sporadic and epidemic strains. Journal of Food Protection 67, 2496–2499.

Brouwer, M.C., van de Beek, D., Heckenberg, S.G.B., Spanjaard, L., de Gans, J., 2006. Community-acquired Listeria monocytogenes meningitis in adults. Clin Infect Dis. 43, 1233–1238. doi:10.1086/508462

Browne, N., Heelan, M., Kavanagh, K., 2013. An analysis of the structural and functional similarities of insect hemocytes and mammalian phagocytes. Virulence 4, 597–603. doi:10.4161/viru.25906

Brynildsrud, O., Bohlin, J., Scheffer, L., Eldholm, V., 2016. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. Genome Biol. 17, 238. doi:10.1186/s13059-016-1108-8

Buchrieser, C., 2007. Biodiversity of the species Listeria monocytogenes and the genus Listeria. Microbes and Infection 9, 1147–1155. doi:10.1016/j.micinf.2007.05.002

Buchrieser, C., Rusniok, C., Kunst, F., Cossart, P., Glaser, P., Listeria Consortium, 2003. Comparison of the genome sequences of Listeria monocytogenes and Listeria innocua: clues for evolution and pathogenicity. FEMS Immunol. Med. Microbiol. 35, 207–213.

Camejo, A., Buchrieser, C., Couvé, E., Carvalho, F., Reis, O., Ferreira, P., Sousa, S., Cossart, P., Cabanes, D., 2009. In Vivo Transcriptional Profiling of Listeria monocytogenes and Mutagenesis Identify New Virulence Factors Involved in Infection. PLoS Pathogens 5, e1000449. doi:10.1371/journal.ppat.1000449

Camejo, A., Carvalho, F., Reis, O., Leitão, E., Sousa, S., Cabanes, D., 2011. The arsenal of virulence factors deployed by Listeria monocytogenes to promote its cell infection cycle. Virulence 2, 379–394. doi:10.4161/viru.2.5.17703

Charlier, C., Perrodeau, É., Leclercq, A., Cazenave, B., Pilmis, B., Henry, B., Lopes, A., Maury, M.M., Moura, A., Goffinet, F., Bracq-Dieye, H., Thouvenot, P., Ungeheuer, M.N., Tourdjman, M., Goulet, V., de Valk, H., Lortholary, O., Ravaud, P., Lecuit, M., 2017. Clinical features and prognostic factors of listeriosis: the MONALISA national prospective cohort study. Lancet Infect Dis 17, 510–519. doi:10.1016/S1473-3099(16)30521-7

Chenal-Francisque, V., Lopez, J., Cantinelli, T., Caro, V., Tran, C., Leclercq, A., Lecuit, M., Brisse, S., 2011. Worldwide distribution of major clones of Listeria monocytogenes. Emerg. Infect. Dis. 17, 1110–1112. doi:10.3201/eid/1706.101778

Cheng, L.W., Portnoy, D.A., 2003. Drosophila S2 cells: an alternative infection model for Listeria monocytogenes. Cell. Microbiol. 5, 875–885.

Chlebicz, A., Śliżewska, K., 2018. Campylobacteriosis, Salmonellosis, Yersiniosis, and Listeriosis as Zoonotic Foodborne Diseases: A Review. IJERPH 15, 863. doi:10.3390/ijerph15050863

Ciesielczuk, H., Betts, J., Phee, L., Doumith, M., Hope, R., Woodford, N., Wareham, D.W., 2015. Comparative virulence of urinary and bloodstream isolates of extra-intestinal pathogenic Escherichia coli in a Galleria mellonella model. Virulence 6, 145–151. doi:10.4161/21505594.2014.988095

Collins, C., Didelot, X., 2018. A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination. PLoS Comput Biol 14, e1005958. doi:10.1371/journal.pcbi.1005958

Cook, S.M., McArthur, J.D., 2013. Developing Galleria mellonella as a model host for human pathogens. Virulence 4, 350–353. doi:10.4161/viru.25240

Cossart, P., Toledo-Arana, A., 2008. Listeria monocytogenes, a unique model in infection biology: an overview. Microbes Infect. 10, 1041–1050. doi:10.1016/j.micinf.2008.07.043

Cotter, P.D., Draper, L.A., Lawton, E.M., Daly, K.M., Groeger, D.S., Casey, P.G., Ross, R.P., Hill, C., 2008. Listeriolysin S, a novel peptide haemolysin associated with a subset of lineage I Listeria monocytogenes. PLoS Pathogens 4, e1000144. doi:10.1371/journal.ppat.1000144

Crim, S.M., Iwamoto, M., Griffin, P.M., Tauxe, R., Huang, J.Y., Gilliss, D., Cronquist, A.B., Cartter, M., Tobin-D'Angelo, M., 2014. Incidence and trends of infection with pathogens transmitted commonly through food—Foodborne Diseases Active Surveillance Network, 10 US sites, 2006-2013. Morbidity and Mortality Weekly Report 63, 328–332.

Dalton, C.B., Merritt, T.D., Unicomb, L.E., Kirk, M.D., Stafford, R.J., Lalor, K., OzFoodNet Working Group, 2011. A national case-control study of risk factors for listeriosis in Australia. Epidemiol. Infect. 139, 437–445. doi:10.1017/S0950268810000944

Datta, A.R., Burall, L.S., 2018. Serotype to genotype: The changing landscape of listeriosis outbreak investigations. Food Microbiology 75, 18–27. doi:10.1016/j.fm.2017.06.013

Datta, A.R., Laksanalamai, P., Solomotis, M., 2013. Recent developments in molecular sub-typing of Listeria monocytogenes. Food Additives & Contaminants: Part A 30, 1437–1445. doi:10.1080/19440049.2012.728722

Dewey-Mattia, D., Kisselburgh, H., Manikonda, K., Silver, R., 2018. Surveillance for foodborne disease outbreaks–United States, 2016: annual report.

Dhama, K., Karthik, K., Tiwari, R., Shabbir, M.Z., Barbuddhe, S., Malik, S.V.S., Singh, R.K., 2015. Listeriosis in animals, its public health significance (food-borne zoonosis) and advances in diagnosis and control: a comprehensive review. Vet Q 35, 211–235. doi:10.1080/01652176.2015.1063023

Didelot, X., Wilson, D.J., 2015. ClonalFrameML: efficient inference of recombination in whole bacterial genomes. PLoS Comput Biol 11, e1004041. doi:10.1371/journal.pcbi.1004041

Disson, O., Lecuit, M., 2012. Targeting of the central nervous system by Listeria monocytogenes. Virulence 3, 213–221. doi:10.4161/viru.19586

Dramsi, S., Biswas, I., Maguin, E., Braun, L., Mastroeni, P., Cossart, P., 1995. Entry of Listeria monocytogenes into hepatocytes requires expression of inIB, a surface protein of the internalin multigene family. Mol. Microbiol. 16, 251–261.

Dreyer, M., Aguilar-Bultet, L., Rupp, S., Guldimann, C., Stephan, R., Schock, A., Otter, A., Schüpbach, G., Brisse, S., Lecuit, M., Frey, J., Oevermann, A., 2016. Listeria monocytogenes sequence type 1 is predominant in ruminant rhombencephalitis. Nature Publishing Group 6, 1–11. doi:10.1038/srep36419

Dreyer, M., Thomann, A., Böttcher, S., Frey, J., Oevermann, A., 2015. Outbreak investigation identifies a single Listeria monocytogenes strain in sheep with different clinical manifestations, soil and water. Veterinary Microbiology 179, 69–75. doi:10.1016/j.vetmic.2015.01.025

Dutilh, B.E., Backus, L., Edwards, R.A., Wels, M., Bayjanov, J.R., van Hijum, S.A.F.T., 2013. Explaining microbial phenotypes on a genomic scale: GWAS for microbes. Briefings in Functional Genomics 12, 366–380. doi:10.1093/bfgp/elt008

Earle, S.G., Wu, C.-H., Charlesworth, J., Stoesser, N., Gordon, N.C., Walker, T.M., Spencer, C.C.A., Iqbal, Z., Clifton, D.A., Hopkins, K.L., Woodford, N., Smith, E.G., Ismail, N., Llewelyn, M.J., Peto, T.E., Crook, D.W., McVean, G., Walker, A.S., Wilson, D.J., 2016. Identifying lineage effects when controlling for population structure improves power in bacterial association studies. Nature Microbiology 1, 16041–9. doi:10.1038/nmicrobiol.2016.41

Elinav, H., Hershko-Klement, A., Valinsky, L., Jaffe, J., Wiseman, A., Shimon, H., Braun, E., Paitan, Y., Block, C., Sorek, R., Nir-Paz, R., Miron, D., Glikman, D., Soboh, S., Nseir, W., Paz, A., Cohen, E., Mendelson, B., Paz, E., Shimoni, Z., Wattad, M., Ravid, M., Keller, N., Rahav, G., Dan, M., Shechner, V., Weinberger, M., Nadir, E., Troshin, T., Riesenberg, K., Tsabari, A., Lachish, T., Halperin, E., Anis, E., Vasiliev, V., Japeth, R., 2014. Pregnancy-Associated Listeriosis: Clinical Characteristics and Geospatial Analysis of a 10-Year Period in Israel. Clin Infect Dis. 59, 953–961. doi:10.1093/cid/ciu504

European Food Safety Authority (EFSA), European Centre for Disease Prevention and Control (ECDC), 2017. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food‐borne outbreaks in 2016. EFSA Journal 15. doi:10.2903/j.efsa.2017.5077

Ewels, P., Magnusson, M., Lundin, S., Käller, M., 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32, 3047–3048. doi:10.1093/bioinformatics/btw354

Falush, D., 2016. Bacterial genomics: Microbial GWAS coming of age. Nature Microbiology 1, 16059–3. doi:10.1038/nmicrobiol.2016.59

Falush, D., Bowden, R., 2006. Genome-wide association mapping in bacteria? Trends in Microbiology 14, 353–355. doi:10.1016/j.tim.2006.06.003

Faralla, C., Rizzuto, G.A., Lowe, D.E., Kim, B., Cooke, C., Shiow, L.R., Bakardjiev, A.I., 2016. InlP, a New Virulence Factor with Strong Placental Tropism. Infect. Immun. 84, 3584–3596. doi:10.1128/IAI.00625-16

Fenlon, D.R., Wilson, J., Donachie, W., 1996. The incidence and level of Listeria monocytogenes contamination of food sources at primary production and initial processing. J. Appl. Bacteriol. 81, 641–650.

Finney, D.J., Stevens, W.L., 1948. A table for the calculation of working probits and weights in probit analysis. Biometrika 35, 191–201.

Flamm, R.K., Hinrichs, D.J., Thomashow, M.F., 1984. Introduction of pAM beta 1 into Listeria monocytogenes by conjugation and homology between native L. monocytogenes plasmids. Infect. Immun. 44, 157–161.

Fox, E.M., Allnutt, T., Bradbury, M.I., Fanning, S., Chandry, P.S., 2016. Comparative Genomics of the Listeria monocytogenes ST204 Subgroup. Front. Microbiol. 7, 1–12. doi:10.3389/fmicb.2016.02057

Fritsch, L., Felten, A., Palma, F., Mariet, J.-F., Radomski, N., Mistou, M.-Y., Augustin, J.-C., Guillier, L., 2019. Insights from genome-wide approaches to identify variants associated to phenotypes at pan-genome scale: Application to L. monocytogenes' ability to grow in cold conditions. International Journal of Food Microbiology 291, 181–188. doi:10.1016/j.ijfoodmicro.2018.11.028

Gahan, C.G.M., Hill, C., 2014. Listeria monocytogenes: survival and adaptation in the gastrointestinal tract. Front Cell Infect Microbiol 4, 9. doi:10.3389/fcimb.2014.00009

Gahan, C.G.M., Hill, C., 2005. Gastrointestinal phase of Listeria monocytogenesinfection. J Appl Microbiol 98, 1345–1353. doi:10.1111/j.1365-2672.2005.02559.x

Gaillard, J.L., Berche, P., Frehel, C., Gouin, E., Cossart, P., 1991. Entry of L. monocytogenes into cells is mediated by internalin, a repeat protein reminiscent of surface antigens from gram-positive cocci. Cell 65, 1127–1141.

Garcia, E., De Paz, M., Rodriguez, J.L., 1996, 1996. Exogenous sources of Listeria contamination in raw ewe's milk. Journal of Food Protection 59, 950–954. doi:10.4315/0362-028X-59.9.950

García, J.A., Micheloud, J.F., Campero, C.M., Morrell, E.L., Odriozola, E.R., Moreira, A.R., 2016. Enteric listeriosis in grazing steers supplemented with spoiled silage. J Vet Diagn Invest 28, 65–69. doi:10.1177/1040638715616658

Gardner, S.N., Slezak, T., Hall, B.G., 2015. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. Bioinformatics 31, 2877–2878. doi:10.1093/bioinformatics/btv271

Goldstein, D.Y., Gialanella, P., Levi, M.H., 2013. Photo quiz: a woman on hemodialysis for end-stage renal disease with fever. Journal of Clinical Microbiology 51, 1355–1655. doi:10.1128/JCM.01099-12

Golnazarian, C.A., Donnelly, C.W., Pintauro, S.J., Howard, D.B., 1989. Comparison of Infectious Dose of Listeria monocytogenes F5817 as Determined for Normal Versus Compromised C57B1/6J Mice. Journal of Food Protection 52, 696–701. doi:10.4315/0362-028X-52.10.696

Goodfellow, M., Sutcliffe, I., Chun, J., 2014. New Approaches to Prokaryotic Systematics. Academic Press.

Gouin, E., Mengaud, J., Cossart, P., 1994. The virulence gene cluster of Listeria monocytogenes is also present in Listeria ivanovii, an animal pathogen, and Listeria seeligeri, a nonpathogenic species. Infect. Immun. 62, 3550–3553.

Gözel, B., Monney, C., Aguilar-Bultet, L., Rupp, S., Frey, J., Oevermann, A., 2019. Hyperinvasiveness of Listeria monocytogenes sequence type 1 is independent of lineage I-specific genes encoding internalin-like proteins. Microbiologyopen 81, e790. doi:10.1002/mbo3.790

Haas, A., Dumbsky, M., Kreft, J., 1992. Listeriolysin genes: complete sequence of ilo from Listeria ivanovii and of lso from Listeria seeligeri. Biochim. Biophys. Acta 1130, 81–84.

Haase, J.K., Didelot, X., Lecuit, M., Korkeala, H., L monocytogenes MLST Study Group, Achtman, M., 2014. The ubiquitous nature of Listeria monocytogenes clones: a large-scale Multilocus Sequence Typing study. Environ Microbiol 16, 405–416. doi:10.1111/1462-2920.12342

Hadfield, J., Croucher, N.J., Goater, R.J., Abudahab, K., Aanensen, D.M., Harris, S.R., 2017. Phandango: an interactive viewer for bacterial population genomics. Bioinformatics 34, 292–293. doi:10.1093/bioinformatics/btx610

Hain, T., Chatterjee, S.S., Ghai, R., Kuenne, C.T., Billion, A., Steinweg, C., Domann, E., Kärst, U., Jänsch, L., Wehland, J., Eisenreich, W., Bacher, A., Joseph, B., Schär, J., Kreft, J., Klumpp, J., Loessner, M.J., Dorscht, J., Neuhaus, K., Fuchs, T.M., Scherer, S., Doumith, M., Jacquet, C., Martin, P., Cossart, P., Rusniock, C., Glaser, P., Buchrieser, C., Goebel, W., Chakraborty, T., 2007. Pathogenomics of Listeria spp. International Journal of Medical Microbiology 297, 541–557. doi:10.1016/j.ijmm.2007.03.016

Hain, T., Ghai, R., Billion, A., Kuenne, C.T., Steinweg, C., Izar, B., Mohamed, W., Mraheil, M.A., Domann, E., Schaffrath, S., Kärst, U., Goesmann, A., Oehm, S., Pühler, A., Merkl, R., Vorwerk, S., Glaser, P., Garrido, P., Rusniok, C., Buchrieser, C., Goebel, W., Chakraborty, T., 2012. Comparative genomics and transcriptomics of lineages I, II, and III strains of Listeria monocytogenes. BMC Genomics 13, 1–1. doi:10.1186/1471-2164-13-144

Haley, B.J., Sonnier, J., Schukken, Y.H., Karns, J.S., Van Kessel, J.A.S., 2015. Diversity of Listeria monocytogenesWithin a U.S. Dairy Herd, 2004–2010. Foodborne Pathogens and Disease 150901135833003–7. doi:10.1089/fpd.2014.1886

Hernandez-Milian, A., Payeras-Cifre, A., 2014. What is new in listeriosis? BioMed Research International 2014, 358051–7. doi:10.1155/2014/358051

Hilliard, A., Leong, D., O'Callaghan, A., Culligan, E.P., Morgan, C.A., DeLappe, N., Hill, C., Jordan, K., Cormican, M., Gahan, C.G.M., 2018. Genomic Characterization of Listeria monocytogenes Isolates Associated with Clinical Listeriosis and the Food Production Environment in Ireland. Genes (Basel) 9, 171. doi:10.3390/genes9030171

Hingston, P., Chen, J., Dhillon, B.K., Laing, C., Bertelli, C., Gannon, V., Tasara, T., Allen, K., Brinkman, F.S.L., Truelstrup Hansen, L., Wang, S., 2017. Genotypes Associated with Listeria monocytogenes Isolates Displaying Impaired or Enhanced Tolerances to Cold, Salt, Acid, or Desiccation Stress. Front. Microbiol. 8, 369. doi:10.3389/fmicb.2017.00369

Hoelzer, K., Pouillot, R., Dennis, S., 2012. Animal models of listeriosis: a comparative review of the current state of the art and lessons learned. Vet. Res. 43, 18. doi:10.1186/1297-9716-43-18

Hoffmann, J.A., 1995. Innate immunity of insects. Curr. Opin. Immunol. 7, 4–10.

Jeffers, G.T., Bruce, J.L., McDonough, P.L., Scarlett, J., Boor, K.J., Wiedmann, M., 2001. Comparative genetic characterization of Listeria monocytogenes isolates from human and animal listeriosis cases. Microbiology 147, 1095–1104. doi:10.1099/00221287-147-5-1095

Joyce, S.A., Gahan, C.G.M., 2010. Molecular pathogenesis of Listeria monocytogenes in the alternative model host Galleria mellonella. Microbiology (Reading, Engl.) 156, 3456–3468. doi:10.1099/mic.0.040782-0

Kale, S.B., Doijad, S.P., Poharkar, K.V., Garg, S., Pathak, A.J., Raorane, A.V., Rawool, D.B., Kurkure, N.V., Barbuddhe, S.B., 2017. Elucidation of the role of the minC gene in filament formation by Listeria monocytogenes under stress conditions. IJCRR 9.

Kathariou, S., Evans, P., Dutta, V., 2017. Strain-Specific Virulence Differences in Listeria monocytogenes: Current Perspectives in Addressing an Old and Vexing Issue, in: Gurtler, J.B., Doyle, M.P., Kornacki, J.L. (Eds.), Foodborne Pathogens. Springer International Publishing, Cham, pp. 61–92. doi:10.1007/978-3-319-56836-2_3

Kavanagh, K., Reeves, E.P., 2004. Exploiting the potential of insects for in vivo pathogenicity testing of microbial pathogens. FEMS Microbiol. Rev. 28, 101–112. doi:10.1016/j.femsre.2003.09.002

Kérouanton, A., Marault, M., Petit, L., Grout, J., Dao, T.T., Brisabois, A., 2010. Evaluation of a multiplex PCR assay as an alternative method for Listeria monocytogenes serotyping. Journal of Microbiological Methods 80, 134–137. doi:10.1016/j.mimet.2009.11.008

Killiny, N., 2018. Generous hosts: Why the larvae of greater wax moth, Galleria mellonella is a perfect infectious host model? Virulence 9, 860–865. doi:10.1080/21505594.2018.1454172

Kim, M., Oh, H.-S., Park, S.-C., Chun, J., 2014. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int. J. Syst. Evol. Microbiol. 64, 346–351. doi:10.1099/ijs.0.059774-0

Kim, S.W., Haendiges, J., Keller, E.N., Myers, R., Kim, A., Lombard, J.E., Karns, J.S., Van Kessel, J.A.S., Haley, B.J., 2018. Genetic diversity and virulence profiles of Listeria monocytogenes recovered from bulk tank milk, milk filters, and milking equipment from dairies in the United States (2002 to 2014). PLoS ONE 13, e0197053. doi:10.1371/journal.pone.0197053

Korsak, D., Krawczyk-Balska, A., 2016. Rifampicin- and Rifabutin-Resistant Listeria monocytogenes Strains Isolated from Food Products Carry Mutations in rpoB Gene. Foodborne Pathogens and Disease 13, 363–368. doi:10.1089/fpd.2015.2103

Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., Stamatakis, A., 2019. RAxML-NG: A fast, scalable, and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 71, 34. doi:10.1093/bioinformatics/btz305

Kremer, P.H.C., Lees, J.A., Koopmans, M.M., Ferwerda, B., Arends, A.W.M., Feller, M.M., Schipper, K., Valls Seron, M., van der Ende, A., Brouwer, M.C., van de Beek, D., Bentley, S.D., 2017. Benzalkonium tolerance genes and outcome in Listeria monocytogenes meningitis. Clin. Microbiol. Infect. 23, 265.e1–265.e7. doi:10.1016/j.cmi.2016.12.008

Kuenne, C., Billion, A., Mraheil, M.A., Strittmatter, A., Daniel, R., Goesmann, A., Barbuddhe, S., Hain, T., Chakraborty, T., 2013. Reassessment of the Listeria monocytogenes pan-genome reveals dynamic integration hotspots and mobile genetic elements as major components of the accessory genome. BMC Genomics 14, 47. doi:10.1186/1471-2164-14-47

Lamond, N.M., Freitag, N.E., 2018. Vertical Transmission of Listeria monocytogenes: Probing the Balance between Protection from Pathogens and Fetal Tolerance. Pathogens 7, 52. doi:10.3390/pathogens7020052

Le Monnier, A., Autret, N., Join-Lambert, O.F., Jaubert, F., Charbit, A., Berche, P., Kayal, S., 2007. ActA is required for crossing of the fetoplacental barrier by Listeria monocytogenes. Infect. Immun. 75, 950–957. doi:10.1128/IAI.01570-06

Lecuit, M., 2005. Understanding how Listeria monocytogenes targets and crosses host barriers. Clinical Microbiology and Infection 11, 430–436. doi:10.1111/j.1469-0691.2005.01146.x

Lee, S., Chen, Y., Gorski, L., Ward, T.J., Osborne, J., Kathariou, S., 2018. Listeria monocytogenes Source Distribution Analysis Indicates Regional Heterogeneity and Ecological Niche Preference among Serotype 4b Clones. MBio 9, e00396–18. doi:10.1128/mBio.00396-18

Lee, S., Ward, T.J., Siletzky, R.M., Kathariou, S., 2012. Two novel type II restriction-modification systems occupying genomically equivalent locations on the chromosomes of Listeria monocytogenes strains. Applied and Environmental Microbiology 78, 2623–2630. doi:10.1128/AEM.07203-11

Letunic, I., Bork, P., 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 44, W242–5. doi:10.1093/nar/gkw290

Levraud, J.-P., Disson, O., Kissa, K., Bonne, I., Cossart, P., Herbomel, P., Lecuit, M., 2009. Real-time observation of listeria monocytogenes-phagocyte interactions in living zebrafish larvae. Infect. Immun. 77, 3651–3660. doi:10.1128/IAI.00408-09

Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22, 1658–1659. doi:10.1093/bioinformatics/btl158

Loh, J.M.S., Adenwalla, N., Wiles, S., Proft, T., 2013. Galleria mellonella larvae as an infection model for group A streptococcus. Virulence 4, 419–428. doi:10.4161/viru.24930

Lomonaco, S., Nucera, D., Filipello, V., 2015. The evolution and epidemiology of Listeria monocytogenes in Europe and the United States. Infection, Genetics and Evolution 35, 172–183. doi:10.1016/j.meegid.2015.08.008

Madjunkov, M., Chaudhry, S., Ito, S., 2017. Listeriosis during pregnancy. Arch. Gynecol. Obstet. 296, 143–152. doi:10.1007/s00404-017-4401-1

Marder, E.P., MPH, Griffin, P.M., Cieslak, P.R., Dunn, J., Hurd, S., Jervis, R., Lathrop, S., Muse, A., Ryan, P., Smith, K., Tobin-D'Angelo, M., Vugia, D.J., Holt, K.G., Wolpert, B.J., Tauxe, R., Geissler, A.L., 2018. Preliminary Incidence and Trends of Infections with Pathogens Transmitted Commonly Through Food — Foodborne Diseases Active Surveillance Network, 10 U.S. Sites, 2006–2017. Morbidity and Mortality Weekly Report 67, 324–328. doi:10.15585/mmwr.mm6711a3

Maury, M.M., Bracq-Dieye, H., Huang, L., Vales, G., Lavina, M., Thouvenot, P., Disson, O., Leclercq, A., Brisse, S., Lecuit, M., 2019. Hypervirulent Listeria monocytogenes clones' adaption to mammalian gut accounts for their association with dairy products. Nat Comms 10, 2488. doi:10.1038/s41467-019-10380-0

Maury, M.M., Tsai, Y.-H., Charlier, C., Touchon, M., Chenal-Francisque, V., Leclercq, A., Criscuolo, A., Gaultier, C., Roussel, S., Brisabois, A., Disson, O., Rocha, E.P.C., Brisse, S., Lecuit, M., 2016a. Uncovering Listeria monocytogenes hypervirulence by harnessing its biodiversity. Nat. Genet. 48, 308–313. doi:10.1038/ng.3501

Maury, M.M., Tsai, Y.-H., Charlier, C., Touchon, M., Chenal-Francisque, V., Leclercq, A., Criscuolo, A., Gaultier, C., Roussel, S., Brisabois, A., Disson, O., Rocha, E.P.C., Brisse, S., Lecuit, M., 2016b. Uncovering Listeria monocytogenes hypervirulence by harnessing its biodiversity. Nat. Genet. 48, 308–313. doi:10.1038/ng.3501

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees, in:. Presented at the Gateway Computing Environments Workshop GCE. doi:10.1109/GCE.2010.5676129","publicationTitle“:”2010

Mook, P., Grant, K.A., Little, C.L., Kafatos, G., Gillespie, I.A., 2010. Emergence of pregnancy-related listeriosis amongst ethnic minorities in England and Wales. Euro Surveill. 15, 17–23.

Moura, A., Disson, O., Lavina, M., Thouvenot, P., Huang, L., Leclercq, A., Fredriksson-Ahomaa, M., Eshwar, A.K., Stephan, R., Lecuit, M., 2019. Atypical Hemolytic Listeria innocua Isolates Are Virulent, albeit Less than Listeria monocytogenes. Infect. Immun. 87, 5273. doi:10.1128/IAI.00758-18

Mukherjee, K., Altincicek, B., Hain, T., Domann, E., Vilcinskas, A., Chakraborty, T., 2010. Galleria mellonella as a model system for studying Listeria pathogenesis. Applied and Environmental Microbiology 76, 310–317. doi:10.1128/AEM.01301-09

Nightingale, K.K., Fortes, E.D., Ho, A.J., Schukken, Y.H., Gröhn, Y.T., Wiedmann, M., 2005. Evaluation of farm management practices as risk factors for clinical listeriosis and fecal shedding of Listeria monocytogenes in ruminants. J. Am. Vet. Med. Assoc. 227, 1808–1814.

Nightingale, K.K., Ivy, R.A., Ho, A.J., Fortes, E.D., Njaa, B.L., Peters, R.M., Wiedmann, M., 2008. inlA premature stop codons are common among Listeria monocytogenes isolates from foods and yield virulence-attenuated strains that confer protection against fully virulent strains. Applied and Environmental Microbiology 74, 6570–6583. doi:10.1128/AEM.00997-08

Nightingale, K.K., Schukken, Y.H., Nightingale, C.R., Fortes, E.D., Ho, A.J., Her, Z., Gröhn, Y.T., McDonough, P.L., Wiedmann, M., 2004. Ecology and transmission of Listeria monocytogenes infecting ruminants and in the farm environment. Applied and Environmental Microbiology 70, 4458–4467. doi:10.1128/AEM.70.8.4458-4467.2004

Nyarko, E.B., Donnelly, C.W., 2015. Listeria monocytogenes: Strain Heterogeneity, Methods, and Challenges of Subtyping. J. Food Sci. 80, M2868–78. doi:10.1111/1750-3841.13133

Oevermann, A., Zurbriggen, A., Vandevelde, M., 2010. Rhombencephalitis Caused by Listeria monocytogenes in Humans and Ruminants: A Zoonosis on the Rise? Interdiscip Perspect Infect Dis 2010, 632513–22. doi:10.1155/2010/632513

Orsi, R.H., Bakker, den, H.C., Wiedmann, M., 2011. Listeria monocytogenes lineages: Genomics, evolution, ecology, and phenotypic characteristics. Int. J. Med. Microbiol. 301, 79–96. doi:10.1016/j.ijmm.2010.05.002

Orsi, R.H., Wiedmann, M., 2016. Characteristics and distribution of Listeria spp., including Listeria species newly described since 2009. Applied Microbiology and Biotechnology 100, 1–15. doi:10.1007/s00253-016-7552-2

Page, A.J., Cummins, C.A., Hunt, M., Wong, V.K., Reuter, S., Holden, M.T.G., Fookes, M., Falush, D., Keane, J.A., Parkhill, J., 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31, 3691–3693. doi:10.1093/bioinformatics/btv421

Pagliano, P., Arslan, F., Ascione, T., 2017. Epidemiology and treatment of the commonest form of listeriosis: meningitis and bacteraemia. Infez Med 25, 210–216.

Pieta, L., Guzman, F., Moya, M.L.E., Campos, F.S., Frazzon, A.P.G., Frazzon, J., 2018. Comparative genomics suggests differences related to resistance and virulence between food-isolated Listeria monocytogenes serotypes 1/2a and 4b.

Pohl, A.M., Pouillot, R., Van Doren, J.M., 2017. Changing US Population Demographics: What Does This Mean for Listeriosis Incidence and Exposure? Foodborne Pathogens and Disease 14, 524–530. doi:10.1089/fpd.2017.2297

Pohl, M.A., Wiedmann, M., Nightingale, K.K., 2006. Associations among Listeria monocytogenes genotypes and distinct clinical manifestations of listeriosis in cattle. Am. J. Vet. Res. 67, 616–626. doi:10.2460/ajvr.67.4.616

Portnoy, D.A., Chakraborty, T., Goebel, W., Cossart, P., 1992. Molecular determinants of Listeria monocytogenes pathogenesis. Infect. Immun. 60, 1263–1267.

Poulin, R., Combes, C., 1999. The concept of virulence: interpretations and implications. Parasitol Today 15, 474–475. doi:10.1016/S0169-4758(99)01554-9

Power, R.A., Parkhill, J., de Oliveira, T., 2017. Microbial genome-wide association studies: lessons from human GWAS. Nature Publishing Group 18, 41–50. doi:10.1038/nrg.2016.132

Quereda, J.J., Meza-Torres, J., Cossart, P., Pizarro-Cerdá, J., 2017a. Listeriolysin S: A bacteriocin from epidemic Listeria monocytogenes strains that targets the gut microbiota. Gut Microbes 8, 384–391. doi:10.1080/19490976.2017.1290759

Quereda, J.J., Nahori, M.A., Meza-Torres, J., Sachse, M., Titos-Jiménez, P., Gomez-Laguna, J., Dussurget, O., Cossart, P., Pizarro-Cerdá, J., 2017b. Listeriolysin S Is a Streptolysin S-Like Virulence Factor That Targets Exclusively Prokaryotic Cells In Vivo. MBio 8, 19484. doi:10.1128/mBio.00259-17

Radoshevich, L., Cossart, P., 2018. Listeria monocytogenes: towards a complete picture of its physiology and pathogenesis. Nat. Rev. Microbiol. 16, 32–46. doi:10.1038/nrmicro.2017.126

Rakic Martinez, M., Wiedmann, M., Ferguson, M., Datta, A.R., 2017. Assessment of Listeria monocytogenes virulence in the Galleria mellonella insect larvae model. PLoS ONE 12, e0184557. doi:10.1371/journal.pone.0184557

Reddy, S., Lawrence, M.L., 2014. Virulence characterization of Listeria monocytogenes. Methods Mol. Biol. 1157, 157–165. doi:10.1007/978-1-4939-0703-8_13

Rocha, C.E., Mol, J.P.S., Garcia, L.N.N., Costa, L.F., Santos, R.L., Paixão, T.A., 2017. Comparative experimental infection of Listeria monocytogenes and Listeria ivanovii in bovine trophoblasts. PLoS ONE 12, e0176911. doi:10.1371/journal.pone.0176911

Rocha, P.R.D., Lomonaco, S., Bottero, M.T., Dalmasso, A., Dondo, A., Grattarola, C., Zuccon, F., Iulini, B., Knabel, S.J., Capucchio, M.T., Casalone, C., 2013. Ruminant rhombencephalitis-associated Listeria monocytogenes strains constitute a genetically homogeneous group related to human outbreak strains. Applied and Environmental Microbiology 79, 3059–3066. doi:10.1128/AEM.00219-13

Rodríguez-López, P., Rodríguez-Herrera, J.J., Vázquez-Sánchez, D., López Cabo, M., 2018. Current Knowledge on Listeria monocytogenes Biofilms in Food-Related Environments: Incidence, Resistance to Biocides, Ecology and Biocontrol. Foods 7, 85. doi:10.3390/foods7060085

Rolhion, N., Cossart, P., 2017. How the study of Listeria monocytogenes has led to new concepts in biology. Future Microbiology 12, 621–638. doi:10.2217/fmb-2016-0221

Roulo, R.M., Fishburn, J.D., Amosu, M., Etchison, A.R., Smith, M.A., 2014. Dose response of Listeria monocytogenes invasion, fetal morbidity, and fetal mortality after oral challenge in pregnant and nonpregnant Mongolian gerbils. Infect. Immun. 82, 4834–4841. doi:10.1128/IAI.01514-14

Rupp, S., Aguilar-Bultet, L., Jagannathan, V., Guldimann, C., Drögemüller, C., Pfarrer, C., Vidondo, B., Seuberlich, T., Frey, J., Oevermann, A., 2015. A naturally occurring prfA truncation in a Listeria monocytogenes field strain contributes to reduced replication and cell-to-cell spread. Veterinary Microbiology 179, 91–101. doi:10.1016/j.vetmic.2015.03.002

Rupp, S., Bärtschi, M., Frey, J., Oevermann, A., 2017. Hyperinvasiveness and increased intercellular spread of Listeria monocytogenes sequence type 1 are independent of listeriolysin S, internalin F and internalin J1. Journal of Medical Microbiology 66, 1053–1062. doi:10.1099/jmm.0.000529

Sauders, B.D., Overdevest, J., Fortes, E., Windham, K., Schukken, Y., Lembo, A., Wiedmann, M., 2012. Diversity of Listeria species in urban and natural environments. Applied and Environmental Microbiology 78, 4420–4433. doi:10.1128/AEM.00282-12

Scalfaro, C., Iacobino, A., Nardis, C., Franciosa, G., 2017. Galleria mellonella as an in vivo model for assessing the protective activity of probiotics against gastrointestinal bacterial pathogens. FEMS Microbiology Letters 364, 2253. doi:10.1093/femsle/fnx064

Scallan, E., Hoekstra, R.M., Angulo, F.J., Tauxe, R.V., Widdowson, M.-A., Roy, S.L., Jones, J.L., Griffin, P.M., 2011. Foodborne illness acquired in the United States--major pathogens. Emerg. Infect. Dis. 17, 7–15. doi:10.3201/eid1701.091101p1

Schmitz-Esser, S., Müller, A., Stessl, B., Wagner, M., 2015. Genomes of sequence type 121 Listeria monocytogenes strains harbor highly conserved plasmids and prophages. Front. Microbiol. 6, 380. doi:10.3389/fmicb.2015.00380

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. doi:10.1093/bioinformatics/btu153

Sheppard, S.K., Didelot, X., Meric, G., Torralbo, A., Jolley, K.A., Kelly, D.J., Bentley, S.D., Maiden, M.C.J., Parkhill, J., Falush, D., 2013. Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in Campylobacter. Proc. Natl. Acad. Sci. U.S.A. 110, 11923–11927. doi:10.1073/pnas.1305559110

Silk, B.J., Date, K.A., Jackson, K.A., Pouillot, R., Holt, K.G., Graves, L.M., Ong, K.L., Hurd, S., Meyer, R., Marcus, R., Shiferaw, B., Norton, D.M., Medus, C., Zansky, S.M., Cronquist, A.B., Henao, O.L., Jones, T.F., Vugia, D.J., Farley, M.M., Mahon, B.E., 2012. Invasive listeriosis in the Foodborne Diseases Active Surveillance Network (FoodNet), 2004-2009: further targeted prevention needed for higher-risk groups. Clin Infect Dis. 54 Suppl 5, S396–404. doi:10.1093/cid/cis268

Smith, G.A., Theriot, J.A., Portnoy, D.A., 1996. The tandem repeat domain in the Listeria monocytogenes ActA protein controls the rate of actin-based motility, the percentage of moving bacteria, and the localization of vasodilator-stimulated phosphoprotein and profilin. J. Cell Biol. 135, 647–660. doi:10.1083/jcb.135.3.647

Steckler, A.J., Cardenas-Alvarez, M.X., Townsend Ramsett, M.K., Dyer, N., Bergholz, T.M., 2018. Genetic characterization of Listeria monocytogenes from ruminant listeriosis from different geographical regions in the U.S. Veterinary Microbiology 215, 93–97. doi:10.1016/j.vetmic.2017.12.021

Stelma, G.N., Reyes, A.L., Peeler, J.T., Francis, D.W., Hunt, J.M., Spaulding, P.L., Johnson, C.H., Lovett, J., 1987. Pathogenicity test for Listeria monocytogenes using immunocompromised mice. Journal of Clinical Microbiology 25, 2085–2089.

Stessl, B., Rückerl, I., Wagner, M., 2014. Multilocus sequence typing (MLST) of Listeria monocytogenes. Methods Mol. Biol. 1157, 73–83. doi:10.1007/978-1-4939-0703-8_6

Suzuki, M., Shibayama, K., Yahara, K., 2016. A genome-wide association study identifies a horizontally transferred bacterial surface adhesin gene associated with antimicrobial resistant strains. Scientific Reports 2016 6 6, 37811. doi:10.1038/srep37811

Šteingolde, Ž., Avsejenko, J., Berzins, A., 2014. Overview of Listeria monocytogenes caused abortions in cattle in Latvia in 2013. Research for Rural Development 1.

Tan, M.F., Siow, C.C., Dutta, A., Mutha, N.V., Wee, W.Y., Heydari, H., Tan, S.Y., Ang, M.Y., Wong, G.J., Choo, S.W., 2015. Development of ListeriaBase and comparative analysis of Listeria monocytogenes. BMC Genomics 16, 755. doi:10.1186/s12864-015-1959-5

Thomsen, L.E., Slutz, S.S., Tan, M.-W., Ingmer, H., 2006. Caenorhabditis elegans is a model host for Listeria monocytogenes. Applied and Environmental Microbiology 72, 1700–1701. doi:10.1128/AEM.72.2.1700-1701.2006

Tinsley, C.R., Bille, E., Nassif, X., 2006. Bacteriophages and pathogenicity: more than just providing a toxin? Microbes Infect. 8, 1365–1371. doi:10.1016/j.micinf.2005.12.013

Vazquez-Boland, J.A., Dominguez, L., Blanco, M., Rocourt, J., Fernández-Garayzábal, J.F., Gutiérrez, C.B., Tascón, R.I., Rodriguez-Ferri, E.F., 1992. Epidemiologic investigation of a silage-associated epizootic of ovine listeric encephalitis, using a new Listeria-selective enumeration medium and phage typing. Am. J. Vet. Res. 53, 368–371.

Vazquez-Boland, J.A., Domínguez-Bernal, G., González-Zorn, B., Kreft, J., Goebel, W., 2001. Pathogenicity islands and virulence evolution in Listeria. Microbes Infect. 3, 571–584.

Vázquez-Boland, J.A., Krypotou, E., Scortti, M., 2017. Listeria Placental Infection. MBio 8, 1125. doi:10.1128/mBio.00949-17

Walland, J., Lauper, J., Frey, J., Imhof, R., Stephan, R., Seuberlich, T., Oevermann, A., 2015. Listeria monocytogenes infection in ruminants: Is there a link to the environment, food and human health? A review. Schweiz. Arch. Tierheilkd. 157, 319–328. doi:10.17236/sat00022

Wand, M.E., McCowen, J.W.I., Nugent, P.G., Sutton, J.M., 2013. Complex interactions of Klebsiella pneumoniae with the host immune system in a Galleria mellonella infection model. Journal of Medical Microbiology 62, 1790–1798. doi:10.1099/jmm.0.063032-0

Wang, Q., Lu, Q., Zhao, H., 2015. A review of study designs and statistical methods for genomic epidemiology studies using next generation sequencing. Front. Genet. 6, 1–12. doi:10.3389/fgene.2015.00149

Wang, S., Orsi, R.H., Tang, S., Zhang, W., Wiedmann, M., Boor, K.J., 2014. Phosphotransferase system-dependent extracellular growth of listeria monocytogenes is regulated by alternative sigma factors σL and σH. Applied and Environmental Microbiology 80, 7673–7682. doi:10.1128/AEM.02530-14

Wernars, K., Heuvelman, K., Notermans, S., Domann, E., Leimeister-Wächter, M., Chakraborty, T., 1992. Suitability of the prfA gene, which encodes a regulator of virulence genes in Listeria monocytogenes, in the identification of pathogenic Listeria spp. Applied and Environmental Microbiology 58, 765–768.

Wesley, L.V., 1999. Listeriosis in animals, in: Ryser, E.T., Marth, E.H. (Eds.), Listeria: Listeriosis, and Food Safety. pp. 39–73.

Wiedmann, M., 2002. Molecular subtyping methods for Listeria monocytogenes. J AOAC Int 85, 524–531.

Wiedmann, M., Bruce, J.L., Keating, C., Johnson, A.E., McDonough, P.L., Batt, C.A., 1997. Ribotypes and virulence gene polymorphisms suggest three distinct Listeria monocytogenes lineages with differences in pathogenic potential. Infect. Immun. 65, 2707–2716.

Wiedmann, M., Czajka, J., Bsat, N., Bodis, M., Smith, M.C., Divers, T.J., Batt, C.A., 1994. Diagnosis and epidemiological association of Listeria monocytogenes strains in two outbreaks of listerial encephalitis in small ruminants. Journal of Clinical Microbiology 32, 991–996.

Willet, C.E., Wade, C.M., 2014. From the Phenotype to the Genotype via Bioinformatics, in: Clinical Bioinformatics, Methods in Molecular Biology. Springer New York, New York, NY, pp. 1–16. doi:10.1007/978-1-4939-0847-9_1

Williams, D., Irvin, E.A., Chmielewski, R.A., Frank, J.F., Smith, M.A., 2007. Dose-response of Listeria monocytogenes after oral exposure in pregnant guinea pigs. Journal of Food Protection 70, 1122–1128.

Wingett, S.W., Andrews, S., 2018. FastQ Screen: A tool for multi-genome mapping and quality control. F1000Res 7, 1338. doi:10.12688/f1000research.15931.2

Wolfe, B., Wiepz, G.J., Schotzko, M., Bondarenko, G.I., Durning, M., Simmons, H.A., Mejia, A., Faith, N.G., Sampene, E., Suresh, M., Kathariou, S., Czuprynski, C.J., Golos, T.G., 2017. Acute Fetal Demise with First Trimester Maternal Infection Resulting from Listeria monocytogenes in a Nonhuman Primate Model. MBio 8, 227. doi:10.1128/mBio.01938-16

Zundel, E., Bernard, S., 2006. Listeria monocytogenes translocates throughout the digestive tract in asymptomatic sheep. Journal of Medical Microbiology 55, 1717–1723. doi:10.1099/jmm.0.46709-0

APPENDIX

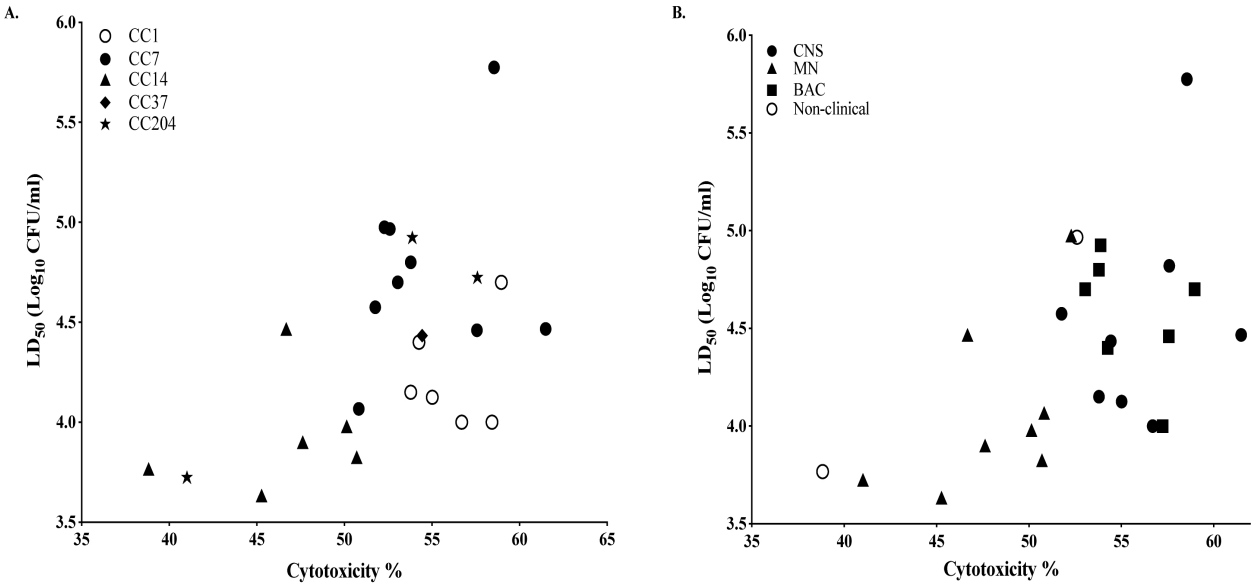

Fig A1. Correlation of LD_{50} and cytotoxicity
Strains grouped (A) by clonal complex (CC) (B) by clinical outcome, ($p<0.001$).

Table A1. Strains used for the genomic assessment of *L. monocytogenes* virulence

Strain ID	Aliases	Source	Clinical Outcome	Lineage	CC	ST	SRR#	Agency/Study source
TB0140	FSL C1-122	Human	BAC	1	1	1	SRR7690645	NYS DoH
TB0435	J4791	Human	BAC	1	1	1	SRR7690673	CDC
TB0548	2014L-6562	Human	BAC	1	1	1	SRR1597488	CDC
TB0549	2014L-6572	Human	BAC	1	1	1	SRR1597496	CDC
TB0550	2014L-6680	Human	BAC	1	1	1	SRR1746766	CDC
TB0551	2014L-6699	Human	BAC	1	1	1	SRR1746767	CDC
TB0552	2014L-6708	Human	BAC	1	1	1	SRR1746768	CDC
TB0096	FSL F2-091	Human	BAC	1	2	2	SRR7690625	NYC DoH
TB0098	FSL F2-661	Human	BAC	1	2	2	SRR7690662	NYS DoH
TB0158	FSL F3-566	Human	BAC	1	2	2	SRR7690659	NYC DoH
TB0427	J3275	Human	BAC	1	2	2	SRR7690639	CDC
TB0565	2016L-6337	Human	BAC	1	2	782	SRR3993242	CDC
TB0532	2013L-5396	Human	BAC	1	4	219	SRR974869	CDC
TB0437	J5119	Human	BAC	1	5	5	SRR7690606	CDC
TB0541	2014L-6264	Human	BAC	1	5	5	SRR1393998	CDC
TB0543	2014L-6391	Human	BAC	1	5	5	SRR1534988	CDC
TB0529	2013L-5283	Human	BAC	1	6	6	SRR945163	CDC
TB0557	2015L-6604	Human	BAC	1	6	6	SRR2968993	CDC
TB0559	2015L-6668	Human	BAC	1	6	6	SRR2480531	CDC
TB0562	2016L-6178	Human	BAC	1	6	6	SRR3510304	CDC
LM07-01337		Human	BAC	1	54	54	ERR1100924	Maury *et al*, 2016
2006/00918	LM06-00918	Human	BAC	1	59	59	ERR1100941	Maury *et al*, 2016
LM08-01516		Human	BAC	1	77	77	ERR1100923	Maury *et al*, 2016
TB0159	FSL F3-757	Human	BAC	1	87	87	SRR7690581	MI DoH
TB0542	2014L-6377	Human	BAC	1	213	213	SRR1534974	CDC
TB0403	FSL R9-0953	Bovine	BAC	1	217	1240	*This study*	Cornell FSL

Table A1. Strains used for the genomic assessment of *L. monocytogenes* virulence (continued)

Strain ID	Aliases	Source	Clinical Outcome	Lineage	CC	ST	SRR#	Agency/Study source
TB0544	2014L-6495	Human	BAC	1	217	217	SRR1575069	CDC
TB0563	2016L-6230	Human	BAC	1	217	217	SRR3723072	CDC
TB0568	2016L-6445	Human	BAC	1	217	217	SRR4228348	CDC
TB0429	J3661	Human	BAC	1	224	224	SRR7690643	CDC
TB0432	J3933	Human	BAC	1	288	288	SRR6214504	CDC
TB0436	J4822	Human	BAC	1	288	288	SRR7690667	CDC
TB0547	2014L-6561	Human	BAC	1	382	382	SRR1597487	CDC
TB0554	2014L-6766	Human	BAC	1	382	382	SRR1746770	CDC
TB0533	2013L-5417	Human	BAC	1	388	388	SRR974878	CDC
TB0556	2015L-6603	Human	BAC	1	554	554	SRR2968992	CDC
TB0567	2016L-6413	Human	BAC	1	554	554	SRR4046805	CDC
TB0569	2016L-6665	Human	BAC	1	554	554	SRR5120729	CDC
TB0553	2014L-6749	Human	BAC	1	554	999	SRR2049025	CDC
TB0535	2013L-5575	Human	BAC	1	639	639	SRR1033764	CDC
TB0443	2009L-1489	Human	BAC	1	663	663	SRR2051159	CDC
TB0633	FSL F2-0375	Human	BAC	1	5	5	*This study*	Cornell FSL
TB0434	J4696	Human	BAC	1	639	639	SRR7690672	CDC
TB0088	FSL F3-631	Human	BAC	2	7	7	SRR7690657	OH DoH
TB0390	FSL J2-0007	Bovine	BAC	2	7	7	*This study*	Cornell FSL
TB0452	4883	Bovine	BAC	2	7	7	*This study*	ND VDL
TB0561	2016L-6098	Human	BAC	2	7	7	SRR3215136	CDC
LM05-00401		Human	BAC	2	8	8	ERR1100940	Maury *et al*, 2016
TB0570	2016L-6776	Human	BAC	2	8	8	SRR5264133	CDC
TB0438	J5492	Human	BAC	2	9	9	SRR6214325	CDC
TB0093	FSL F2-141	Human	BAC	2	11	11	SRR7690669	NYC DoH
TB0094	FSL F2-405	Human	BAC	2	11	11	SRR7690666	NYS DoH

Table A1. Strains used for the genomic assessment of *L. monocytogenes* virulence (continued)

Strain ID	Aliases	Source	Clinical Outcome	Lineage	CC	ST	SRR#	Agency/Study source
TB0138	FSL C1-111	Human	BAC	2	11	11	SRR7690629	NYS DoH
TB0431	J3877	Human	BAC	2	11	11	SRR6214442	CDC
LM07-01377		Human	BAC	2	14	14	ERR1100946	Maury *et al*, 2016
TB0566	2016L-6338	Human	BAC	2	19	378	SRR3993244	CDC
LM06-01244		Human	BAC	2	20	20	ERR1100928	Maury *et al*, 2016
TB0449	2011L-2614	Human	BAC	2	29	29	SRR6212584	CDC
TB0531	2013L-5361	Human	BAC	2	29	29	SRR972390	CDC
LM09-01099		Human	BAC	2	37	37	ERR1100926	Maury *et al*, 2016
LM07-01398		Human	BAC	2	121	121	ERR1100947	Maury *et al*, 2016
TB0546	2014L-6523	Human	BAC	2	121	121	SRR1575091	CDC
TB0137	FSL C1-051	Human	BAC	2	155	155	SRR7690630	OH DoH
TB0448	2011L-2436	Human	BAC	2	155	155	SRR6214469	CDC
TB0564	2016L-6312	Human	BAC	2	155	155	SRR3826948	CDC
TB0361	15-5477	Calf	BAC	2	199	230	*This study*	ND VDL
TB0393	FSL J2-0018	Bovine	BAC	2	204	204	*This study*	Cornell FSL
TB0089	FSL F3-744	Human	BAC	2	376	376	SRR7690646	MI DoH
TB0560	2015L-6871	Human	BAC	2	1055	1055	SRR3028136	CDC
TB0440	2009L-1297	Human	BAC	2	1076	1076	SRR6214499	CDC
TB0439	2009L-1236	Human	BAC	2	1367	1367	SRR6214500	CDC
TB0430	J3767	Human	BAC	3	450	1051	SRR7690656	CDC
LM05-00008		Human	CNS	1	1	1	ERR1100936	Maury *et al*, 2016
LM05-00172		Human	CNS	1	1	1	ERR1100938	Maury *et al*, 2016
LM07-00596		Human	CNS	1	1	1	ERR1100943	Maury *et al*, 2016
LM05-00190	LMNC284	Ruminant	CNS	1	1	1	ERR1590242	Dreyer *et al*, 2016
LM07-01398	LMNC320	Ruminant	CNS	1	1	1	ERR1590246	Dreyer *et al*, 2016
LM07-01377	LMNC479	Cattle	CNS	1	1	1	ERR1590247	Dreyer *et al*, 2016

Table A1. Strains used for the genomic assessment of *L. monocytogenes* virulence (continued)

Strain ID	Aliases	Source	Clinical Outcome	Lineage	CC	ST	SRR#	Agency/Study source
LM07-01398	LMNC480	Cattle	CNS	1	1	1	ERR1590248	Dreyer *et al*, 2016
LM05-00190	LMNC048	Cattle	CNS	1	1	1	ERR1590249	Dreyer *et al*, 2016
LM05-00401	LMNC050	Cattle	CNS	1	1	1	ERR1590252	Dreyer *et al*, 2016
LM09-00261	LMNC482	Cattle	CNS	1	1	1	ERR1590253	Dreyer *et al*, 2016
LM09-00372	LMNC483	Goat	CNS	1	1	1	ERR1590254	Dreyer *et al*, 2016
LM09-00558	LMNC484	Sheep	CNS	1	1	1	ERR1590258	Dreyer *et al*, 2016
LM09-00865	LMNC485	Sheep	CNS	1	1	1	ERR1590259	Dreyer *et al*, 2016
LM09-01004	LMNC486	Cattle	CNS	1	1	1	ERR1590261	Dreyer *et al*, 2016
LM09-01092	LMNC487	Cattle	CNS	1	1	1	ERR1590262	Dreyer *et al*, 2016
LM05-00065	LMNC023	Cattle	CNS	1	1	1	ERR1590263	Dreyer *et al*, 2016
JF5052	LMNC01	Cattle	CNS	1	1	1	ERR1590264	Dreyer *et al*, 2016
LM05-00172	LMNC025	Cattle	CNS	1	1	1	ERR1590266	Dreyer *et al*, 2016
LM10-00508	LMNC489	Cattle	CNS	1	1	1	ERR1590267	Dreyer *et al*, 2016
LM06-01465	LMNC079	Goat	CNS	1	1	1	ERR1590269	Dreyer *et al*, 2016
LM06-01686	LMNC081	Goat	CNS	1	1	1	ERR1590270	Dreyer *et al*, 2016
LM05-00065	LMNC281	Ruminant	CNS	1	1	1	ERR1590273	Dreyer *et al*, 2016
LM09-00865	LMNC328	Ruminant	CNS	1	1	1	ERR1590274	Dreyer *et al*, 2016
LM10-00507	LMNC337	Ruminant	CNS	1	1	1	ERR1590277	Dreyer *et al*, 2016
LM06-00584	LMNC378	Goat	CNS	1	1	1	ERR1590280	Dreyer *et al*, 2016
LM06-02349	LMNC380	Goat	CNS	1	1	1	ERR1590281	Dreyer *et al*, 2016
LM06-02570	LMNC381	Goat	CNS	1	1	1	ERR1590282	Dreyer *et al*, 2016
LM07-00596	LMNC382	Sheep	CNS	1	1	1	ERR1590283	Dreyer *et al*, 2016
TB0359	15-5128	Bovine	CNS	1	1	1	*This study*	ND VDL
TB0509	543	Sheep	CNS	1	1	1	*This study*	ND VDL
TB0510	650	Sheep	CNS	1	1	1	*This study*	ND VDL
TB0513	1031-1	Sheep	CNS	1	1	1	*This study*	ND VDL

Table A1. Strains used for the genomic assessment of *L. monocytogenes* virulence (continued)

Strain ID	Aliases	Source	Clinical Outcome	Lineage	CC	ST	SRR#	Agency/Study source
TB0514	1031-2	Sheep	CNS	1	1	1	*This study*	ND VDL
TB0545	2014L-6496	Human	CNS	1	1	1	SRR1575070	CDC
LM08-01519	LMNC481	Cattle	CNS	1	1	413	ERR1590250	Dreyer *et al*, 2016
LM06-00582	LMNC058	Sheep	CNS	1	1	414	ERR1590255	Dreyer *et al*, 2016
LM09-01099	LMNC488	Cattle	CNS	1	1	515	ERR1590265	Dreyer *et al*, 2016
TB0396	FSL J2-0022	Bovine	CNS	1	1	1239	*This study*	Cornell FSL
LM07-00876		Human	CNS	1	3	3	ERR1100935	Maury *et al*, 2016
LM09-01092	LMNC265	Ruminant	CNS	1	4	4	ERR1590238	Dreyer *et al*, 2016
LM95800	LMNC271	Sheep	CNS	1	4	4	ERR1590239	Dreyer *et al*, 2016
LM06-01907	LMNC302	Ruminant	CNS	1	4	4	ERR1590243	Dreyer *et al*, 2016
LM06-02128	LMNC303	Ruminant	CNS	1	4	4	ERR1590244	Dreyer *et al*, 2016
LM07-01067	LMNC316	Ruminant	CNS	1	4	4	ERR1590245	Dreyer *et al*, 2016
LM07-01337	LMNC385	Cattle	CNS	1	4	4	ERR1590251	Dreyer *et al*, 2016
LM05-00008	LMNC278	Ruminant	CNS	1	4	4	ERR1590272	Dreyer *et al*, 2016
LM09-01004	LMNC331	Ruminant	CNS	1	4	4	ERR1590275	Dreyer *et al*, 2016
LM09-01099	LMNC336	Ruminant	CNS	1	4	4	ERR1590276	Dreyer *et al*, 2016
LM07-00876	LMNC383	Sheep	CNS	1	4	4	ERR1590284	Dreyer *et al*, 2016
LM07-01067	LMNC384	Sheep	CNS	1	4	4	ERR1590285	Dreyer *et al*, 2016
TB0621	569	Bovine	CNS	1	4	219	SRR7690607	ND VDL
LM09-00372		Human	CNS	1	6	6	ERR1100913	Maury *et al*, 2016
LM06-01023		Human	CNS	1	6	6	ERR1100942	Maury *et al*, 2016
LM09-01092		Human	CNS	1	6	6	ERR1100951	Maury *et al*, 2016
JF5052	LMNC277	Ruminant	CNS	1	6	6	ERR1590271	Dreyer *et al*, 2016
LM05-00008	LMNC357	Ruminant	CNS	1	6	6	ERR1590278	Dreyer *et al*, 2016
LM05-00065	LMNC358	Ruminant	CNS	1	6	6	ERR1590279	Dreyer *et al*, 2016
TB0528	2013L-5275	Human	CNS	1	6	6	SRR945156	CDC

Table A1. Strains used for the genomic assessment of *L. monocytogenes* virulence (continued)

Strain ID	Aliases	Source	Clinical Outcome	Lineage	CC	ST	SRR#	Agency/Study source
TB0558	2013L-5195	Human	CNS	1	6	6	SRR2422614	CDC
LM06-00581		Human	CNS	1	6	615	ERR1100954	Maury *et al*, 2016
JF4978		Ruminant	CNS	1	59	59	ERR1590257	Dreyer *et al*, 2016
JF4942	LMNC060	Sheep	CNS	1	59	59	ERR1590268	Dreyer *et al*, 2016
TB0404	JF4982	Sheep	CNS	1	191	191	*This study*	ND VDL
TB0405	JF5052	Sheep	CNS	1	191	191	*This study*	ND VDL
LM10-00505		Human	CNS	1	315	194	ERR1100930	Maury *et al*, 2016
TB0579	7358	Bovine	CNS	1	379	379	*This study*	ND VDL
TB0364	15-7302	Bovine	CNS	1	554	554	*This study*	SDSU
JF4982	LMNC274	Ruminant	CNS	1	663	663	ERR1590241	Dreyer *et al*, 2016
LM09-00261		Human	CNS	2	7	7	ERR1100949	Maury *et al*, 2016
LM10-00506	LMNC269	Ruminant	CNS	2	7	7	ERR1600115	Dreyer *et al*, 2016
TB0358	15-4516	Bovine	CNS	2	7	7	*This study*	ND VDL
TB0360	15-5428	Bovine	CNS	2	7	7	*This study*	ND VDL
TB0451	4269	Bovine	CNS	2	7	1399	*This study*	ND VDL
LM07-01377	LMNC319	Ruminant	CNS	2	8	16	ERR1600126	Dreyer *et al*, 2016
LM09-00558	LMNC327	Ruminant	CNS	2	8	16	ERR1600129	Dreyer *et al*, 2016
LM07-01345	LMNC462	Sheep	CNS	2	8	16	ERR1600134	Dreyer *et al*, 2016
LM09-01004		Human	CNS	2	9	9	ERR1100950	Maury *et al*, 2016
LM07-00596	LMNC307	Ruminant	CNS	2	11	11	ERR1600123	Dreyer *et al*, 2016
LM07-00596	LMNC084	Sheep	CNS	2	11	11	ERR1600130	Dreyer *et al*, 2016
JF4982	LMNC351	Ruminant	CNS	2	11	11	ERR1600144	Dreyer *et al*, 2016
LM08-01517	LMNC244	Ruminant	CNS	2	14	399	ERR1600109	Dreyer *et al*, 2016
LM09-00261	LMNC247	Ruminant	CNS	2	14	399	ERR1600110	Dreyer *et al*, 2016
LM09-00558	LMNC253	Ruminant	CNS	2	14	399	ERR1600111	Dreyer *et al*, 2016
LM09-00865	LMNC257	Ruminant	CNS	2	14	399	ERR1600112	Dreyer *et al*, 2016

Table A1. Strains used for the genomic assessment of *L. monocytogenes* virulence (continued)

Strain ID	Aliases	Source	Clinical Outcome	Lineage	CC	ST	SRR#	Agency/Study source
JF4978	LMNC340	Ruminant	CNS	2	14	399	ERR1600143	Dreyer *et al*, 2016
LMNC301	LMNC301	Ruminant	CNS	2	18	18	ERR1600122	Dreyer *et al*, 2016
LM07-00876	LMNC314	Ruminant	CNS	2	18	18	ERR1600124	Dreyer *et al*, 2016
JF4982		Ruminant	CNS	2	18	18	ERR1600131	Dreyer *et al*, 2016
JF4982	LM83088	Human	CNS	2	19	19	ERR1100952	Maury *et al*, 2016
TB0428	J3540	Human	CNS	2	20	1508	SRR7690644	CDC
TB0363	15-5911	Sheep	CNS	2	21	21	*This study*	ND VDL
LM44807	LMNC270	Ruminant	CNS	2	29	29	ERR1600116	Dreyer *et al*, 2016
JF5052	LMNC354	Ruminant	CNS	2	29	29	ERR1600145	Dreyer *et al*, 2016
LM05-00401	LMNC365	Ruminant	CNS	2	29	29	ERR1600148	Dreyer *et al*, 2016
LM05-00172	LMNC283	Ruminant	CNS	2	36	36	ERR1600118	Dreyer *et al*, 2016
LM07-01377	LMNC236	Ruminant	CNS	2	37	37	ERR1600107	Dreyer *et al*, 2016
LM07-01398	LMNC241	Ruminant	CNS	2	37	37	ERR1600108	Dreyer *et al*, 2016
LM09-01004	LMNC264	Ruminant	CNS	2	37	37	ERR1600113	Dreyer *et al*, 2016
LM09-01099	LMNC266	Ruminant	CNS	2	37	37	ERR1600114	Dreyer *et al*, 2016
LM07-01337	LMNC317	Ruminant	CNS	2	37	37	ERR1600125	Dreyer *et al*, 2016
LM146793	LMNC338	Ruminant	CNS	2	37	37	ERR1600141	Dreyer *et al*, 2016
LMNC379		Goat	CNS	2	37	37	ERR1600150	Dreyer *et al*, 2016
TB0392	FSL J2-0016	Bovine	CNS	2	37	37	*This study*	Cornell FSL
TB0388	FSL J2-0002	Bovine	CNS	2	89	391	*This study*	Cornell FSL
TB0512	1116	Bovine	CNS	2	121	121	*This study*	ND VDL
TB0381	FSL E1-0003	Bovine	CNS	2	121	1217	SRR7690648	Cornell FSL
TB0362	15-5475	Cow	CNS	2	199	230	*This study*	ND VDL
LM08-01518	LMNC323	Ruminant	CNS	2	200	200	ERR1600127	Dreyer *et al*, 2016
LM09-00261	LMNC325	Ruminant	CNS	2	204	204	ERR1600128	Dreyer *et al*, 2016
TB0485	6804	Goat	CNS	2	204	204	*This study*	ND VDL

Table A1. Strains used for the genomic assessment of *L. monocytogenes* virulence (continued)

Strain ID	Aliases	Source	Clinical Outcome	Lineage	CC	ST	SRR#	Agency/Study source
LM197786	LMNC339	Ruminant	CNS	2	226	226	ERR1600142	Dreyer *et al*, 2016
LM05-00704	LMNC371	Ruminant	CNS	2	226	226	ERR1600149	Dreyer *et al*, 2016
TB0453	16-8643	Bovine	CNS	2	262	1058	*This study*	SDSU
JF5052		Ruminant	CNS	2	412	412	ERR1600132	Dreyer *et al*, 2016
LM05-00008	LMNC022	Sheep	CNS	2	412	412	ERR1600135	Dreyer *et al*, 2016
LM05-00704	LMNC055	Goat	CNS	2	412	412	ERR1600137	Dreyer *et al*, 2016
LM07-00876	LMNC088	Ruminant	CNS	2	412	412	ERR1600138	Dreyer *et al*, 2016
LM05-00190	LMNC361	Ruminant	CNS	2	412	412	ERR1600147	Dreyer *et al*, 2016
LM07-01345		Human	CNS	2	412	620	ERR1100925	Maury *et al*, 2016
LM09-00372	LMNC250	Ruminant	CNS	2	412	1159	ERR1600139	Dreyer *et al*, 2016
JF4978	LMNC273	Goat	CNS	2	415	394	ERR1600117	Dreyer *et al*, 2016
LM06-00583	LMNC297	Ruminant	CNS	2	415	394	ERR1600121	Dreyer *et al*, 2016
LM05-00401	LMNC287	Ruminant	CNS	2	451	451	ERR1600119	Dreyer *et al*, 2016
LM05-00704	LMNC288	Ruminant	CNS	2	451	451	ERR1600120	Dreyer *et al*, 2016
LM09-01092	LMNC332	Ruminant	CNS	2	451	451	ERR1600140	Dreyer *et al*, 2016
LM05-00172	LMNC359	Ruminant	CNS	2	451	451	ERR1600146	Dreyer *et al*, 2016
TB0356	15-3925	Bovine	CNS	2	451	451	*This study*	ND VDL
TB0481	5727	Sheep	CNS	2	451	451	*This study*	ND VDL
TB0534	2013L-5574	Human	CNS	2	451	451	SRR1033763	CDC
TB0571	5251	Bovine	CNS	2	451	451	*This study*	ND VDL
LM07-01345	LMNC318	Ruminant	CNS	3	70	70	ERR1600151	Dreyer *et al*, 2016
LM09-00372	LMNC326	Ruminant	CNS	3	70	70	ERR1600152	Dreyer *et al*, 2016
TB0357	15-4085	Bovine	CNS	3		897	*This study*	ND VDL
TB0402	FSL J2-0069	Bovine	CNS	3	396	1023	*This study*	Cornell FSL
LM07-01337	LMNC163	Cattle	MN	1	1	1	ERR1590260	Dreyer *et al*, 2016
LM09-00865		Human	MN	1	2	2	ERR1100909	Maury *et al*, 2016

Table A1. Strains used for the genomic assessment of *L. monocytogenes* virulence (continued)

Strain ID	Aliases	Source	Clinical Outcome	Lineage	CC	ST	SRR#	Agency/Study source
TB0386	FSL F3-0695	Bovine	MN	1	4	4	*This study*	Cornell FSL
LM09-00558		Human	MN	1	4	4	ERR1100911	Maury *et al*, 2016
LM05-00190		Human	MN	1	4	4	ERR1100939	Maury *et al*, 2016
LM05-00704		Human	MN	1	5	5	ERR1100912	Maury *et al*, 2016
TB0110	FSL J2-064	Bovine	MN	1	5	5	SRR7690617	CU VDL
LM07-01067	LMNC145	Cattle	MN	1	220	220	ERR1590256	Dreyer *et al*, 2016
LM07-01067		Human	MN	1	224	386	ERR1100944	Maury *et al*, 2016
TB0394	FSL J2-0019	Bovine	MN	2	7	7	*This study*	Cornell FSL
LM57179	LM57179	Human	MN	2	9	9	ERR1100972	Maury *et al*, 2016
LM6186	LM6186	Human	MN	2	9	9	ERR1100973	Maury *et al*, 2016
TB0353	15-1274	Bovine	MN	2	14	91	*This study*	ND VDL
TB0354	15-1275	Bovine	MN	2	14	91	*This study*	ND VDL
TB0407	LM05-00008	Bovine	MN	2	14	91	*This study*	ND VDL
TB0408	LM05-00065	Bovine	MN	2	14	91	*This study*	ND VDL
TB0527	4839	Bovine	MN	2	14	91	*This study*	ND VDL
LM07-01345	LMNC164	Cattle	MN	2	18	18	ERR1600136	Dreyer *et al*, 2016
TB0522	3562	Bovine	MN	2	37	37	*This study*	ND VDL
TB0632	5888	Bovine	MN	2	37	37	SRR7690615	ND VDL
TB0577	7301	Bovine	MN	2	90	425	*This study*	ND VDL
JF4978	LM73068	Human	MN	2	121	121	ERR1100974	Maury *et al*, 2016
TB0511	1092	Bovine	MN	2	121	121	*This study*	ND VDL
LM05-00065		Human	MN	2	155	155	ERR1100937	Maury *et al*, 2016
TB0387	FSL F3-0526	Bovine	MN	2	412	412	SRR7690649	Cornell FSL
TB0508	491	Lamb	MN	3	1283	1283	*This study*	ND VDL

*Clinical outcomes: MN: Materno-fetal; CNS: Central Nervous System; BAC: Bacteremia **Agency/Study source: DoH: Department of Health; US states are designated by 2 letters codes; US cities are designated by 3 letters codes; SDSU: South Dakota State University; FSL: Food Safety Laboratory; CDC: Centers for Disease Control and Prevention

Table A2. Virulence genes screened in the *L. monocytogenes* dataset

Gene	Symbol/Aliases	Gene product	Accession number/Location	Pathogenicity Island
actA	lmo0204	actin-assembly inducing protein precursor	NC_003210.1:209470-211389	LIPI-1
hly	lmo0202	listeriolysin O precursor	NC_003210.1:205819-207408	
mpl	lmo0203	Zinc metalloproteinase precursor	NC_003210.1:207739-209271	
inlA	lmo0433	internalin A	NC_003210.1:454534-456936	
inlB	lmo0434	internalin B	NC_003210.1:457021-458913	
plcA	lmo0201	phosphatidylinositol-specific phospholipase c	NC_003210.1:c205577-204624	
plcB	lmo0205	phospholipase C	NC_003210.1:211425-212294	
prfA	lmo0200	listeriolysin positive regulatory protein	NC_003210.1:c204353-203640	
LMOf2365_1113	LMOf2365_1113	ABC transporter, ATP-binding protein	NC_002973.6:1126911-1127798	LIPI-3
LMOf2365_1114	LMOf2365_1114	ABC transporter permease	NC_002973.6:1127798-1128547	
LMOf2365_1115	LMOf2365_1115	hypothetical protein	NC_002973.6:1128588-1128905	
LMOf2365_1116	LMOf2365_1116	SagB/ThcOx family dehydrogenase	NC_002973.6:1128902-1129777	

Table A2. Virulence genes screened in the *L. monocytogenes* dataset (continued)

Gene	Symbol/Aliases	Gene product	Accession number/Location	Pathogenicity Island
LMOf2365_1117	LMOf2365_1117	hypothetical protein	NC_002973.6:1129785-1130702	
LMOf2365_1118	LMOf2365_1118	streptolysin associated protein SagD	NC_002973.6:1130695-1132008	
Lm4b-02327	LM4B_RS11760	PTS lactose/cellobiose transporter subunit IIA	NC_012488.1:2396192-2396506	LIPI-4
Lm4b-02328	LM4B_RS11765	PTS sugar transporter subunit IIB	NC_012488.1:2396506-2396862	
Lm4b-02329	LM4B_RS11770	PTS sugar transporter subunit IIC	NC_012488.1:2396882-2398183	
lmo0444	lmo0444	hypothetical protein	NC_003210.1:473936-476716	
lmo0445	lmo0445	124ranscriptional regulator	NC_003210.1:476960-478447	
lmo0446	lmo0446	penicillin acylase	NC_003210.1:478721-479710	SSI-1
lmo0447	lmo0447	glutamate decarboxylase	NC_003210.1:479765-481153	
lmo0448	lmo0448	amino acid antiporter	NC_003210.1:481250-482701	
ami	lmo2558	autolysin, amidase	NC_003210.1:2635167-2637920	
aut	lmo1076	autolysin	NC_003210.1:1106041-1107759	

Table A2. Virulence genes screened in the *L. monocytogenes* dataset (continued)

Gene	Symbol/Aliases	Gene product	Accession number/Location	Pathogenicity Island
ctaP	lmo0135	peptide ABC transporter substrate-binding protein	NC_003210.1:137323-138897	
fbpA	lmo1829	fibronectin-binding proteins	NC_003210.1:1904152-1905864	
flaA	lmo0690	flagellin	NC_003210.1:724896-725759	
fri	lmo0943	non-heme iron-binding ferritin	NC_003210.1:979059-979529	
gap	lmo2459	glyceraldehyde-3-phosphate dehydrogenase	NC_003210.1:c2532320-2531310	
gtcA	lmo2549	wall teichoic acid glycosylation protein GtcA	NC_003210.1:c2625417-2624980	
hfq	lmo1295	host factor-1 protein	NC_003210.1:1323450-132368	
inlC	lmo1786	internalin C	AL591981.1:107152-108104	
inlF	lmo0409	Internalin F	NC_003210.1:429630-432095	
inlH	lmo0263	internalin H	NC_003210.1:284365-286011	
inlJ	lmo2821	internalin J	NC_003210.1:2907153-2909708	
inlP	lmo2470	Internalin P	NC_003210.1:2544267-2545433	

Table A2. Virulence genes screened in the *L. monocytogenes* dataset (continued)

Gene	Symbol/Aliases	Gene product	Accession number/Location	Pathogenicity Island
lap	lmo1634	bifunctional acetaldehyde-CoA/alcohol	NC_003210.1:1677409-1680009	
lapB	lmo1666	peptidoglycan-linked protein	NC_003210.1:1717193-1722328	
lgt	lmo2482	prolipoprotein diacylglyceryl transferase	NC_003210.1:c2558044-2557211	
lntA	lmo0438	hypothetical protein (LntA)	NC_003210.1:467519-468136	
lpeA	lmo1847	metal ABC transporter	NC_003210.1:1923825-1924757	
lplA1	lmo0931	lipoate protein ligase A	NC_003210.1:967784-968779	
lspA	lmo1844	lipoprotein signal peptidase	NC_003210.1:1920436-1920900	
mprF	lmo1695	hypothetical protein (MprF)	NC_003210.1:1758694-1761291	
murA	lmo2691	autolysin	NC_003210.1:2766935-2768707	
oppA	lmo2196	peptide ABC transporter substrate-binding protein	NC_003210.1:2284539-2286215	
pgdA	lmo0415	endo-1,4-beta-xylanase	NC_003210.1:437482-438882	
pgl	lmo0558	hypothetical protein (Pgl)	NC_003210.1:596580-597620	

Table A2. Virulence genes screened in the *L. monocytogenes* dataset (continued)

Gene	Symbol/Aliases	Gene product	Accession number/Location	Pathogenicity Island
prsA2	lmo2219	foldase	NC_003210.1:2306833-2307714	
pycA	lmo1072	pyruvate carboxylase	NC_003210.1:1099266-1102706	
recA	lmo1398	recombinase A	NC_003210.1:1425419-1426465	
relA	lmo1523	(p)ppGpp synthetase	NC_003210.1:c1558364-1556148	
secA	lmo2510	preprotein translocase subunit SecA	NC_003210.1:c2589813-2587300	
sigB	lmo0895	RNA polymerase sigma factor SigB	NC_003210.1:930671-931450	
sipZ	lmo1271	type I signal peptidase	NC_003210.1:1294360-1294902	
sod	lmo1439	superoxide dismutase	NC_003210.1:c1473588-1472980	
srtA	lmo0929	sortase	NC_003210.1:966245-966913	
svpA	lmo2185	hypothetical protein (SipZ)	NC_003210.1:2272403-2274112	
uhpT	lmo0838	sugar:phosphate antiporter	NC_003210.1:869095-	
vip	lmo0320	peptidoglycan-bound surface protein	NC_003210.1:344850-346049	

Table A2. Virulence genes screened in the *L. monocytogenes* dataset (continued)

Gene	Symbol/Aliases	Gene product	Accession number/Location	Pathogenicity Island
virR	lmo1745	two-component response regulator	NC_003210.1:1814403-1815080	

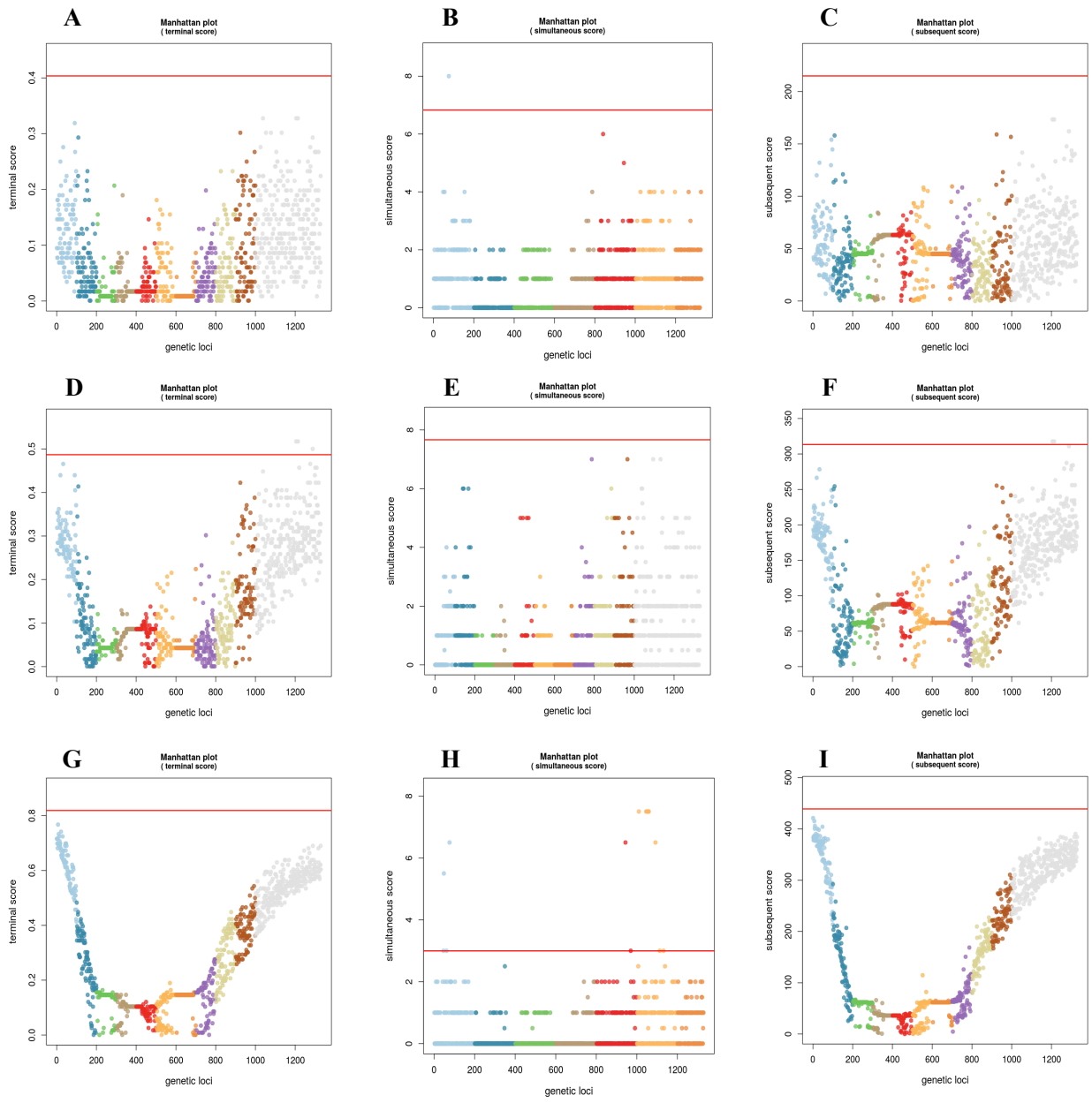

Fig A2. Manhattan plots from the gene-based GWAS using treeWAS
(A, D, G) Score 1, (B, E, H) Score 2, and (C, F, I) Score 3 showing association score values for CNS (A, B, C), BAC (D, E, F), and MN (G, H, I) and all accessory genes. Dots above significant threshold (red line) indicate significant associations ($p<0.05$).

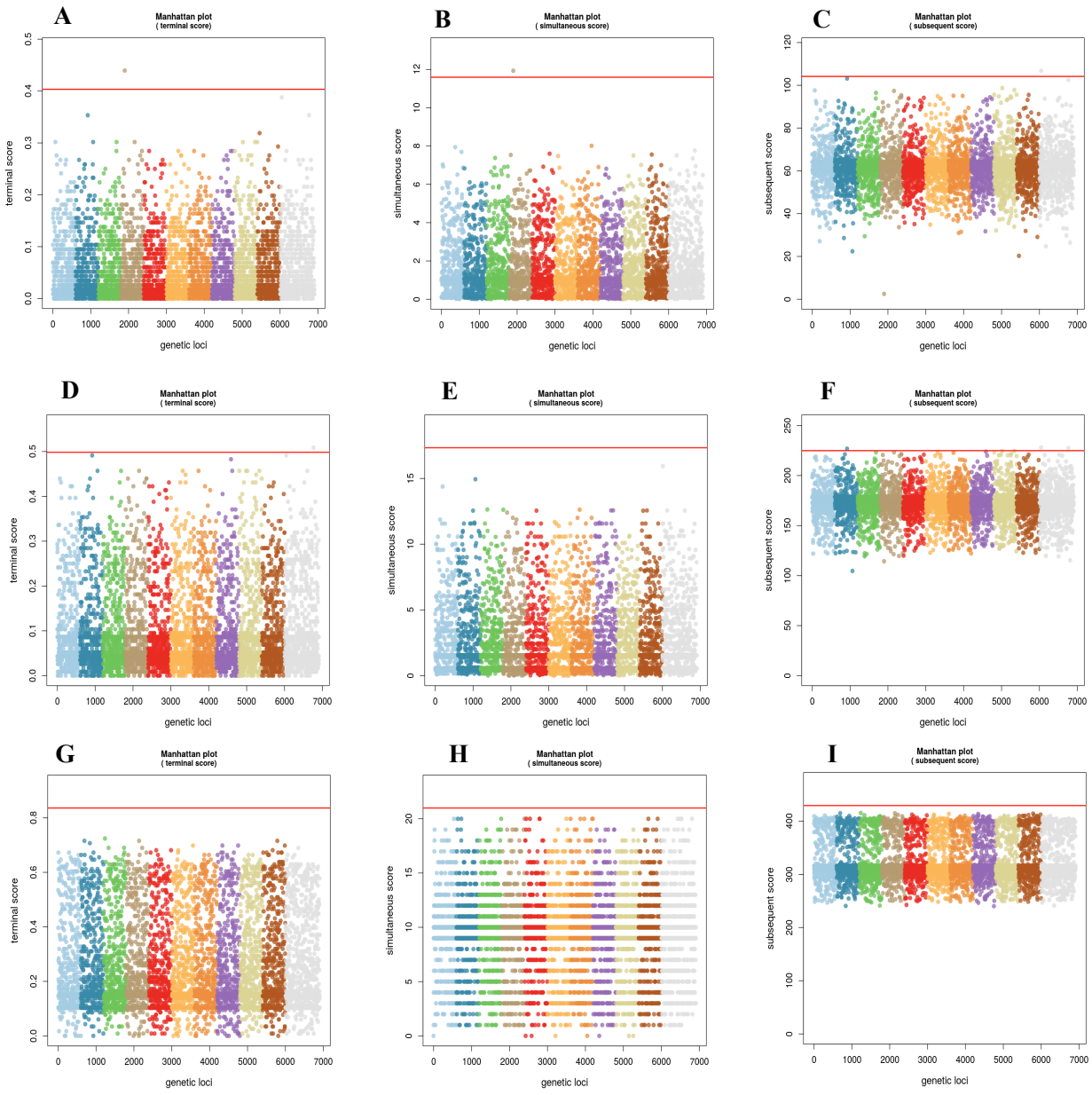

Fig A3. Manhattan plots from the SNP-based GWAS using treeWAS
(A, D, G) Score 1, (B, E, H) Score 2, and (C, F, I) Score 3 showing association score values for CNS (A, B, C), BAC (D, E, F), and MN (G, H, I) and all core-SNPs. Dots above significant threshold (red line) indicate significant associations ($p<0.05$).